

**Interferometrikus
gravitációshullám-detektorok adatának
analízise idő-frekvencia térben**

BSc Szakdolgozat

Angler Gábor

**Témavezető:
Raffai Péter**

Eötvös Loránd Tudományegyetem, Atomfizikai Tanszék

2010. május 28.

Tartalomjegyzék

1. Bevezetés	7
2. A gravitációs sugárzás	9
2.1. A gravitációs hullámokról	9
2.2. Gravitációshullám-detektorok	10
2.3. Asztrofizikai források	12
3. Adatanalízis	15
3.1. Előfeldolgozó eljárások	15
3.1.1. A bemeneti adatsor	15
3.1.2. Adatszűrés	16
3.2. Keresőeljárások	19
3.2.1. Bázisválasztás	19
3.2.2. Koherens és inkoherens módszerek	20
3.2.3. Az Omega Pipeline	21
3.3. Koincidencia analízis	26
4. Képfeldolgozó eljárások	29
4.1. Motiváció	29
4.2. Az algoritmus	31
4.2.1. Keresztkorreláció	31
4.2.2. Az iterációs eljárás	35
4.3. Tesztek	39
4.4. Szimuláció	45
5. Összefoglalás	51

Ábrák jegyzéke

2.1. Az interferometrikus detektorok felépítésének egyszerűsített sémája	11
3.1. Az interferometrikus detektorok érzékenységének irányfüggése	17
3.2. A LIGO detektorainak torzításban kifejezett zajspektrum sűrűsége	18
3.3. A jelenlegi és jövőbeni detektorok spektrális érzékenységgörbéi	19
3.4. Normált Gauss-modulált szinusz bázisfüggvény idő térben	24
3.5. A Q -transzformáció speciális bázisa	26
4.1. Szignifikáns esemény konstans Q érték melletti idő-frekvencia spektrogramjai	32
4.2. Az Omega Pipeline keresőeljárás által rögzített eventgram	38
4.3. Komplex alakzatok összehasonlítása	40
4.4. Alakzatfelismerés során a hasonlóság alakulása az iterációs szám függvényében	41
4.5. Összesített teszteredmények	42
4.6. Alakzatfelismerési tesztekhez használt képek	43
4.7. LIGO detektor adatából előállított eseménytérkép	46
4.8. Események hasonlóságának eloszlás különböző η értékek mellett	50

1. fejezet

Bevezetés

Az einsteini általános relativitáselmélet már évtizedekkel ezelőtt megjósolta egy, a gravitációs kölcsönhatásból származtatható sugárzás létezését. Az ún. gravitációs hullámok fizikai hatásait azóta közvetett módon, csillagászati megfigyelésekkel is sikerült kimutatni. A hullámok közvetlen detektálása csupán napjaink kutatási technológiája mellett vált reálisan elérhető céllá. A közvetlen megfigyelésük nem csak az einsteini elméletet erősítené meg, de az asztrofizikai megfigyelések számára is forradalmian új, az elektromágneses spektrumtól független információt biztosítana. Ez számos asztrofizikai objektum és folyamat, vagy épp a korai Univerzum alaposabb megértését tenné lehetővé.

Napjainkban a Föld több pontján is működnek interferometrikus elven alapuló gravitációshullám-detektorok. Ezek közül a legnagyobb érzékenységgel az Egyesült Államokban található LIGO detektorok bírnak. A LIGO Scientific Collaboration (LSC) a LIGO detektorok fejlesztését, az adatok feldolgozását és asztrofizikai alkalmazását végző szakemberek nemzetközi tudományos közösségét foglalja magába. Az ELTÉ-n 2007-ben megalakult Eötvös Gravity Research Group (EGRG) hazánkból elsőként lett tagja e nemzetközi együttműködésnek, és kapcsolódott be ezzel a gravitációs hullámok kísérleti-elméleti kutatásába.

A kutatási területeink közé tartozik a nagyenergiájú ütközési és kitörési folyamatokból származó, ún. gravitációshullám-kitörések keresése, vizsgálata. Ilyen kitörések forrásai közé tartoznak például a szupernóva-robbanások, vagy a gammasugár-kitörések.

Az ezekből származó jelek várhatóan a néhány tized–száz másodperces időtartományt fedik le, a hullámformák részletes modellje pedig nem ismert. Ez olyan adatfeldolgozási eljárások kidolgozását teszi szükségessé, amelyek a hullámformák széles skáláján képesek robusztus működésre. A diplomamunka során egy ilyen, robusztus képfeldolgozási eljárás kifejlesztését végeztük el, amelyet a gravitációshullám-detektorok adatának idő-frekvencia térbeli reprezentációján alkalmazunk kitörések keresésére.

A fejlesztőmunkánk szervesen kapcsolódik az LSC által használt keresőeljárásokhoz és azok optimalizációjához, munkánkat ezért külföldi kutatópartnerekkel együttműködve végezzük.

2. fejezet

A gravitációs sugárzás

2.1. A gravitációs hullámokról

A csillagászati megfigyelések döntő többsége egészen napjainkig az elektromágneses (EM) sugárzás detektálásán alapult. A XX–XXI. századra a fizikai elméletek pontosabbá válása és a felgyorsult technológiai fejlődés lehetővé tette, hogy a gravitációs kölcsönhatást – mint információhordozó közeget – is vizsgáljuk, akár csillagászati és extragalaktikus méretekben.

Az einsteini általános relativitáselmélet már a XX. század elején megjósolta egy, az elektromágneses hullámokhoz sok tekintetben hasonló, ám a gravitációs kölcsönhatásból származtatható sugárzás létezését. Az elmélet szerint az ún. gravitációs hullámok (*Gravitational Wave*, GW) a téridő metrikájának forrásról leválni képes, fénysebességgel haladó deformációi, melyeket időben gyorsuló tömeg-kvadrupól momentumok keltenek. Az ilyen források lehetnek például szupernóva-robbanások, neutroncsillag és/vagy fekete lyuk ütközések, akár aszimmetrikus neutroncsillagok.

A fizikai hatásukat tekintve a GW-k képesek a téridő azon pontjainak távolságát megváltoztatni, amelyen áthaladnak. Ezen hatás az, amelyet a létezésük közvetlen kimutatására felhasználhatunk. A közvetlen detektálásuk ugyan a mai napig várat magára, jóllehet a létezésükre vonatkozó közvetett bizonyítékok felmutatásáért fizikai Nobel-díjat¹ is kiosztottak már. A GW-k kimutatása azért is bír különleges jelentőséggel, mert a vizsgálatukkal egy, az elektromágneses spektrumtól teljesen független, újabb információhordozó közeg kerülne a birtokunkba. Ez nemcsak az egyébként legkevésbé ismert és megértett gravitációs kölcsönhatás természetére vonatkozó elméletek pontosítását tenné lehetővé, de az asztrofizikának és a koz-

¹ Hulse és Taylor a PSR 1913+16 pulzár-kettősrendszer periódusidejének mérésével megmutatták, hogy a rendszer pályája olyan mértékben zsugorodik, amely kiváló egyezésben áll azon elméleti előrejelzésekkel, amelyek a rendszer összenergiájának csökkenését GW-k emittálásában jelölik meg[1, 2].

mológiának is értékes adatokkal szolgál. Így a gravitációs hullámok kísérleti vizsgálatára az EM-mérések legígéretesebb kiegészítőjeként, s ezzel a jövő asztrofizikájának legfontosabb, ezidáig kiaknázatlan információszerzési lehetőségeként tekinthetünk.

A GW-k az EM-hullámokkal szemben rendkívül gyengén szóródnak vagy nyelődnek el anyagban, így olyan katasztrofikus események belső, dinamikai folyamatairól is információt szerezhetnénk, mint például szupernóvák mag-összeomlása, vagy kompakt kettősrendszerek ütközései. A GW-k detektálása olyan megfigyelési lehetőséget nyújtana, amely elsőként lenne közvetlenül érzékeny az anyag dinamikai mozgására. A feltételezett GW-háttérsugárzás sűrűségfluktuációinak feltérképezése lehetővé tenné a korai univerzum tulajdonságainak vizsgálatát, amely a mai kozmológiai elméleteket vetné a fenntarthatóság próbája alá.

Megjegyezzük, hogy a gravitációs hullámok elméletének részletes ismertetése nem célja ezen dolgozatnak. Ezért a továbbiakban a detektálási folyamat megértéséhez szükséges általános ismeretekre fókuszálunk. A GW-k elméletét nagyobb részletességgel a [8, 23] hivatkozások alatti írások tárgyalják.

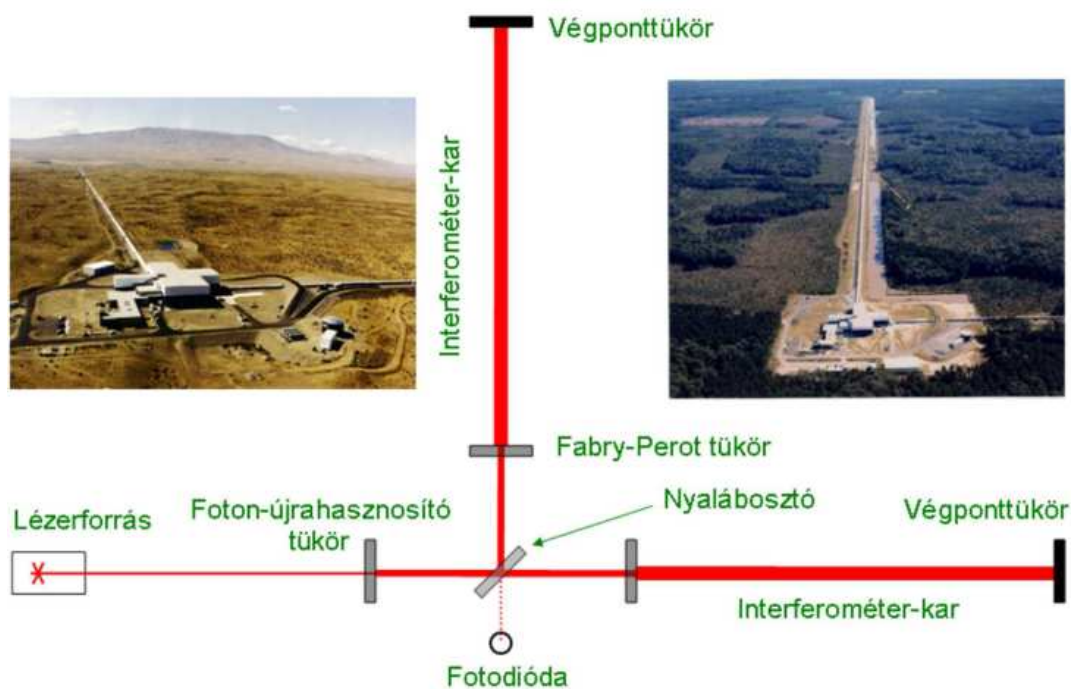
2.2. Gravitációshullám-detektorok

A GW-k közvetlen detektálására a Föld számos pontján épültek detektorok. Ezek hatalmas Michelson-interferométerek², melyek a GW-k téridő torzító hatásait teszik több lépésben észlelhetővé. A működési elvük minden esetben ugyanaz: egy kettéosztott koherens lézernyalábot interferáltatunk önmagával olyan módon, hogy a nyalábok találkozásakor kioltási interferencia valósuljon meg. Egy fotódióddával az ezen pontbeli fényintenzitást regisztráljuk adott mintavételi frekvenciával. Amennyiben a berendezésen GW-jel halad át, az okozott téridő torzulás miatti igen kismértékű optikai úthosszváltozás folytán a kioltási interferencia megszűnik, mely esemény az adatrögzítés során kapott digitális jelsorozatban kimutatható.

A detektálási folyamat technikai kivitelezése igen összetett feladat. Az elméleti jóslatok szerint ugyanis a GW-k a téridőben relatív elmozdulásokat képesek okozni (esetünkben az interferométer végponttükrei között, az eredeti karhosszúsághoz képest), méghozzá igen kis mértékben³.

² Megjegyezzük, hogy Weber úttörő munkásságának köszönhetően már az 1960-as évek óta készítettek ún. tömegrezonátor-detektorokat is (például a *MiniGrail*[5]). Az interferométerek GW-kutatásban történő felhasználási lehetőségeit elsőként R. Weiss tanulmányozta az 1970-es években[3].

³ Ez a gravitációs kölcsönhatás igen gyenge jellegében is megnyilvánul: a GW-k anyaggal csak rendkívül gyengén hatnak kölcsön; lásd a G csatolási tényező értékét. A GW-k okozta torzítás nagyságrendileg 10^{-21} – 10^{-22} [9].



2.1. ábra. Az interferometrikus detektorok felépítésének egyszerűsített sémája, és a LIGO hanfordi (balra) és livingstoni (jobbra) detektorépületei. A lézernyalábot a féligáteresztő tükör kettéválasztja, és a 2, illetve 4 km-re lévő végponttükörök felé továbbítja. A végponttükörök és az ún. Fabry-Perot tükörök közötti többszöri visszaverődés révén az interferométer effektív karhosszúsága megsokszorozható. A foton-újrahasznosító tükör célja a fotonvesztés csökkentése. Az interferenciaképet a fotodióda mintavételezi. (Forrás: [23])

Egy adott GW-jel által okozott

$$h = \frac{\Delta l}{l} \quad (2.1)$$

megnyúlás (vagy torzítás, *strain*) esetén, hogy a távolságváltozás Δl abszolút értékét mérhető nagyságúvá tegyük, szükséges, hogy az interferométer l karhosszát a technikai korlátok figyelembevétele mellett maximalizáljuk.

Az egyesült államokbeli Laser Interferometer Gravitational-Wave Observatory (LIGO) kutatóintézet két 4 km (Hanford, WA, rövidítve: H1; Livingston, LA, rövidítve: L1), valamint egy 2 km (Hanford, WA, rövidítve: H2) (2.1. ábra) karhosszúságú interferométert használ[4]. Ezen GW-antennák⁴ méretüket tekintve jelenleg nemcsak a világ más tájain hasonló céllal létesített detektorokat (VIRGO, Olaszország; GEO600, Németország; TAMA300, Japán[16])

⁴ Megjegyezzük, hogy az „antenna” elnevezés is használatos a GW-detektorokra vonatkozóan. Ugyanis a detektoraink – hasonlóan az EM-antennákhoz – közel minden irányból érkező hullámokat képesek detektálni, ellentétben például egy teleszkóppal, mellyel csak adott égi pozíciók célzott megfigyelését végezhetjük.

haladják meg, de minden idők legérzékenyebb⁵ interferométereinek is számítanak. Az abszolút távolságváltozás, melyet jelenleg a LIGO detektoraival elméletileg még ki lehet mérni $\approx 10^{-18}$ m[9], mely nagyságrendileg egy proton átmérőjének ezredrésze.

A detektálható jelek rendkívül kicsiny volta és a detektorok körüli lokális zajhatások – frekvenciatartománytól, a jel érkezési irányától és polarizációjától függően – igen kicsiny jel-zaj arányt (*SNR*) okozhatnak. Így nemcsak a berendezések zajszintjét kell különféle technológiai megoldásokkal csökkenteni, de a kapott adatsorok elemzésére is minél érzékenyebb feldolgozó eljárásokat kell kifejleszteni. Ezek mérnököknek és fizikusoknak egyaránt igen komoly technikai kihívást jelentenek. A felmerülő problémák nem utolsó sorban kutatócsoportunknak is feladatot biztosítanak.

2.3. Asztrofizikai források

Az előbbieken utaltunk rá, hogy gravitációs hullámokat minden gyorsuló tömeg-kvadrupól-momentummal rendelkező fizikai rendszer képes létrehozni. Igaz, a viszonylag kis tömegű és méretű rendszerek (például Nap-Föld) keltette GW-k deformáló hatása gyakorlati szempontból észlelhetetlen⁶ a jelenlegi detektálási technika mellett. Az elméleti jóslatok alapján a földi, vagy naprendszerbeli detektálható forrás kizárható. Érdeemes megvizsgálni azon ma ismert asztrofizikai forrástípusokat, amelyek által generált GW-jeleket a jelenlegi földi⁷ detektorokkal közvetlenül kimutathatónak tartanak[9, 10].

⁵ Megjegyezzük, hogy a LIGO detektorainak érzékenysége ugrásszerű, jelészlelési valószínűséget[6] tekintve mintegy 15^3 -szoros javulást ér majd el $\approx 2012 - 2014$ -ben, amikor a már kidolgozott tervek alapján és rendelkezésre álló anyagi keret felhasználásával a detektorok továbbfejlesztésre kerülnek (Advanced LIGO Project[7])

⁶ Ennek oka az ilyen esetben kibocsátott GW-k extrém alacsony frekvenciája (a detektorok lokális szeizmikus zaja miatt nem érzékelhetjük) és az igen közeli forrásokból származó GW-k által okozott hatás nem különíthető el a lokális newtoni tér dinamikai változásából eredőtől (ún. gravitációs gradiens zaj).

⁷ Megjegyezzük, hogy az űrbe is terveznek GW-detektort telepíteni: ez a LISA (*Laser Interferometer Space Antenna*), mely három, egymástól 5 millió km-re haladó műholdból állna. Az elméleti várakozások szerint segítségével olyan forrásokból (például szupermasszív fekete lyukaktól) származó – jellemzően igen alacsony frekvenciájú – GW-jeleket is detektálhatnánk, amelyek jelenleg a földi detektorokkal észlelhetetlenek a szeizmikus zaj miatt. A felbocsátása legkorábban 2020-ra várható.

- **Kompakt kettősrendszerek**[1] Ezek a közös tömegközéppontjuk körül keringő neutroncsillag-neutroncsillag, neutroncsillag-fekete lyuk, fekete lyuk-fekete lyuk kettősrendszerek. A mozgásuk során folyamatosan kibocsátanak GW-kat, így energiát veszítenek, pályájuk mérete csökken, miközben keringési frekvenciájuk növekszik. A folyamat során a GW-k frekvenciája és amplitúdója egyaránt nő, így a jel potenciálisan detektálhatóvá válhat a mai földi GW-antennák számára. Az objektumok összeütközése során intenzív GW-kibocsátás várható.

A kompakt kettősrendszerekből származó sugárzás időben rövid és jól modellezett GW-hullámnak tekinthető.

- **GW-háttérsugárzás**[9] A GW-hátteret egyrészt a mikrohullámú EM-háttérhez hasonló, kozmológiai eredetű, feltételezett, sztochasztikus tulajdonságú kozmikus sugárzás, másrészt nagyszámú, fel nem bontható GW-forrás együttes háttérjáruléka alkotja. A kozmológiai eredetű háttérsugárzás forrásai a korai univerzum erős téridő metrikai fluktuációi lehetnek. Ezen sugárzás feltételezetten igen alacsony frekvenciájú, így észlelése elsősorban az űrbe telepítendő detektoroktól várható.

A GW-háttérsugárzás időben hosszú és nem jól modellezett hullám.

- **Periodikus források** A forgómozgást végző aszimmetrikus neutroncsillagot, kettősrendszer ütközés utáni maradványát, kettősrendszerben anyagot intenzíven elnyelő neutroncsillagot kvázi-monokromatikus GW-forrásnak tekintjük.

A periodikus forrásokból származó sugárzás frekvenciája időben jó közelítéssel állandó (kvázi-monokromatikus hullámok), továbbá a hullámforma jól modellezett.

- **Kataklizmikus dinamikai folyamatok** Ilyen folyamatok közé soroljuk nagytömegű csillagok mag-összeomlásait[11], a szupernóva-robbanásokat[12], gammasugár-kitöréseket[13] és bármilyen egyéb nagyenergiájú kitörési folyamatot. Az ilyen folyamatok jelentik becslések szerint a legtöbb, ezidáig ismeretlen jelforrást.

Az feltételezett intenzív GW-emisszió révén a ma működő detektorokkal is várhatóan észlelhetők. A sugárzás anyaggal való gyenge kölcsönhatása miatt a GW-jel korábbi, viszont időben viszonylag közeli detektálását várjuk az ugyanazon forrásból érkező EM és neutrínójelekhez képest. Az egymásnak kölcsönösen megfeleltethető EM/neutrínó és GW-jelek révén a detektorok adatsoraiban célzott időtartományra korlátozható jelkeresésre nyílik lehetőség. Elméleti számítások alapján a jelek maximális hossza becsülhető.

Viszont a különböző dinamikai folyamatok komplexitása miatt a keltett GW-jelek paramétereire vonatkozóan igen nehéz előrejelzéseket tenni. Ezért az ilyen jelek kereséséhez sokkal általánosabb algoritmusok kifejlesztése szükséges, mint például a jól modellezett bespiráló kettősök esetében.

A komplex folyamatokra pontos elméleti modellek kidolgozása igen nagy kihívás az elméleti szakembereknek, amelyet jelentősen megkönnyítheti a kísérleti ellenőrzése már létező, viszont részben hiányos modellek jóslatainak (például felső becslések meghatározása bizonyos modellek szabad paramétereire). Tehát az ilyen jelek detektálása, illetve az azért folytatott erőfeszítések révén közelebb kerülhetünk máig nem tisztázott asztrofizikai folyamatok megértéséhez.

Az ilyen nem-modellezett folyamatokból származó, viszonylag **rövid időskálájú**, viszont **széles frekvenciatartalmú** jeleket röviden **GW-kitöréseknek** (*burst*) nevezük.

3. fejezet

Adatanalízis

A mai gravitációshullám-kutatás egyik meghatározó iránya a nagyenergiájú ütközési és kitörési (*burst*) folyamatokból származó ún. gravitációshullám-kitörések vizsgálata. Ilyen kitörések forrásai például szupernóva-robbanások és gamma-sugár-kitörések (*Gamma Ray Burst*, GRB). Az ezen folyamatokból származó jelek elméleti számítások alapján a néhány tized – száz másodperces időtartományt fedik le, továbbá a hullámformák részletes modellje nem ismert. A detektáláshoz ezért olyan adatfeldolgozási eljárások kifejlesztése szükséges, amelyek hullámformák széles tartományán képesek robusztus működésre. A továbbiakban ismertetésre kerülő, széles körben használt algoritmusok alkalmazásával a gravitációs hullámokat az interferometrikus detektorok jelének idő-frekvencia térbeli reprezentációjában keressük.

Célom elsősorban az, hogy a jelkeresésünk elméleti hátterét, a használt eszközöket bemutassam. Ezért a következőkben csak az algoritmusunk működésének megértéséhez szükséges általános ismeretekre összpontosítunk. A téma egy részletes összefoglalóját találhatja az érdeklődő Olvasó a [23] hivatkozás alatt.

3.1. Előfeldolgozó eljárások

3.1.1. A bemeneti adatsor

A különböző adatfeldolgozó eljárások (ún. *pipeline*-ok) bemeneteként szolgáló kalibrált idő-amplitúdó adatsorokat a tényleges jelkeresésre előkészítő (előfeldolgozó) eljárás sok esetben hasonló, ennek vázlatos tárgyalását a következőkben végezzük.

A rögzített bemeneti adatsor véges hosszúságú, az egyes amplitúdóértékeknek megfelelő számértékek valós időben az adott mintavételi frekvencia (LIGO detektorok esetében: $F_s = 16384$ Hz által definiált diszkrét időlépések szerint követik egymást. Az ideális esetben

végtelenhez tartó hosszúságú bemeneti adatsorra egyúttal egy diszkrét pontokban értelmezett $x(t) : \mathbb{R} \rightarrow \mathbb{R}$ függvényként is tekinthetünk. Tegyük fel, hogy a detektálási idő alatt GW-jelet is észlelünk; ekkor az az adatsorban lineárisan¹ szuperponálódik a folyamatosan jelen lévő zajsorozathoz. Ezért a fentebb definiált $x(t)$ valós függvény két tag összegeként felírható:

$$x(t) = n(t) + h(t), \quad (3.1)$$

ahol $n(t)$ a zaj járulékanak, a $h(t)$ a GW-jel járulékanak megfeleltetett valós függvény.

A zajfüggvény lehet mesterségesen szimulált (leggyakrabban normális eloszlású álvéletlen számok generálásával előállított; keresőalgoritmusok tesztelése céljából) vagy valós zajsorozat (általában nem gaussi eloszlású [23]). A valódi jel mindig leírható két különböző (ortogonális) polarizációjú komponens lineáris kombinációjaként:

$$h(t) = F_{\times} h_{\times}(t) + F_{+} h_{+}(t), \quad (3.2)$$

ahol $h_{\times}(t)$ és $h_{+}(t)$ a jel kétféle polarizációjának megfelelő idő-amplitúdó függvény, az F_{\times} és F_{+} az előbbiekhöz tartozó lineáris együtthatók (ún. antennafaktorok; a detektorok irány- és polarizáció-érzékenységétől függnek) (3.1. ábra).

További részletek az interferometrikus detektorok antennafaktoraira vonatkozóan a [23] hivatkozás alatt található.

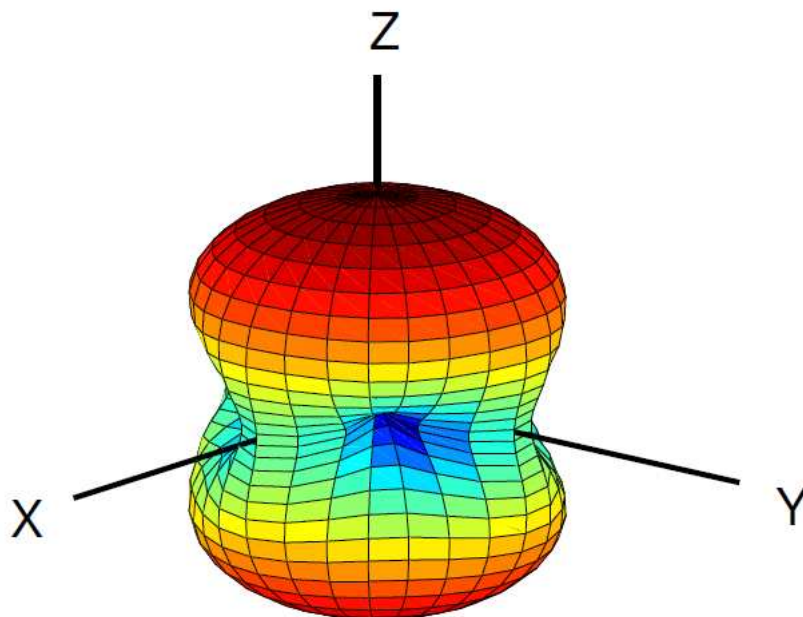
3.1.2. Adatszűrés

A interferometrikus detektorok zaja az érzékelhető jel nagyságot erősen korlátozza. A zajszint erősen frekvenciafüggő, jellemzésére kézenfekvő bevezetni amplitúdóspektrumot. Az amplitúdóspektrum-sűrűséget (*amplitude spectral density*, ASD) lásd a 3.2. ábrán. A jelenlegi és jövőbeni detektorok tervezett érzékenységgörbéit lásd a 3.3. ábrán.

Látható, hogy a zajszint jelentősen magasabb (az érzékenység meredeken leesik) a $\approx 80 - 100$ Hz-nél alacsonyabb, illetve a ≈ 2 kHz-nél magasabb frekvenciatartományokban. Továbbá az érzékenység-görbén megfigyelhetünk számos vékony zajcsúcsot². Kézenfekvőnek tűnik, hogy a rögzített adatsor $\approx 80 - 2000$ Hz intervallumon kívül eső részét sáváteresztő digitális szűrővel (például *Infinite Impulse Response* (IIR) szűrővel) kiszűrjük és a zajcsúcsokat eltávolítsuk (a sáváteresztő szűrő inverzével, *Notch Filtering*), mivel azok jól ismert zajhatások [14] eredményei.

¹ a gyenge tér közelítés feltételeinek teljesülése esetén

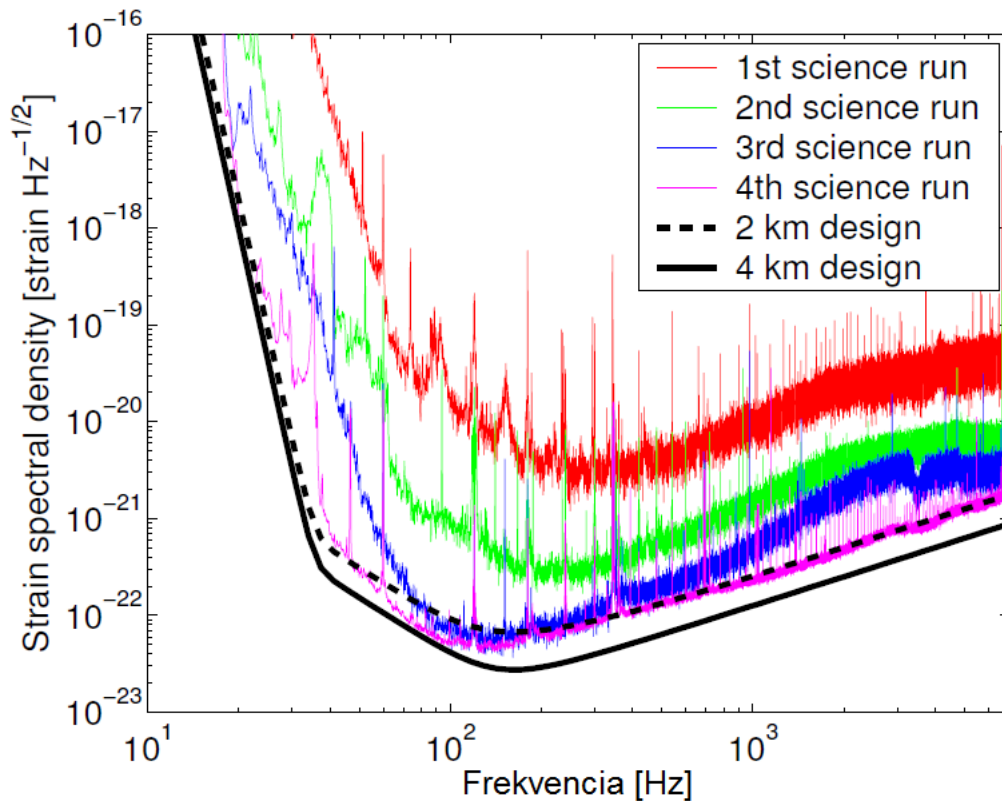
²A csúcsok egy részét a 60 Hz-es hálózati váltakozófeszültség és annak felharmonikusai okozzák. Három további csúcs mesterségesen előidézett, az interferométer kimeneti adatának kalibrálásához szükségesek.



3.1. ábra. Az interferometrikus detektorok érzékenységének irányfüggése véletlenszerű polarizációjú GW-hullámokra átlagolva (az antennafaktorok négyzetösszegének gyökét ábrázoltuk). Az X és Y tengely egybeesik az interferométer karokkal. Látható, hogy az interferometrikus detektorok a Z tengely mentén beeső hullámokra a legérzékenyebbek, míg a síkjukban, a karokkal $\phi = 45^\circ$ szöget bezáró irányból érkező hullámokra érzéketlenek. (Forrás: [8])

A fentebb definiált frekvenciaablak komplementerének eltávolítása idő térben konvolúció alkalmazása, míg Fourier-térben az integrálás szorzási művelet végrehajtására egyszerűsödik. Az alkalmazott előfeldolgozó algoritmusok a digitális szűrést az időtérben hajtják végre, annak ellenére, hogy az számításigényesebb a frekvenciatérbeli szűréshez képest. Ha a szűrést a frekvenciatérben végeznénk, a digitális adatsor diszkrét Fourier-transzformációja (*Discrete Fourier Transform*, DFT) során a vékony zajcsúcsok által képviselt energiaérték egyenletesen oszlana el az elemi frekvenciatartományban, amelyhez tartoznak (*frequency bin*; szélességük általában 2–3 nagyságrenddel nagyobb a csúcsok karakterisztikus szélességénél). Ezen frekvenciabin eltávolítása a vékony csúcsok széles frekvenciatartományát torzítaná. A számításigényesebb időtérbeli konvolúcióval viszont a vékony zajcsúcsok környezetük jelentős torzítása nélkül kiemelhetők.

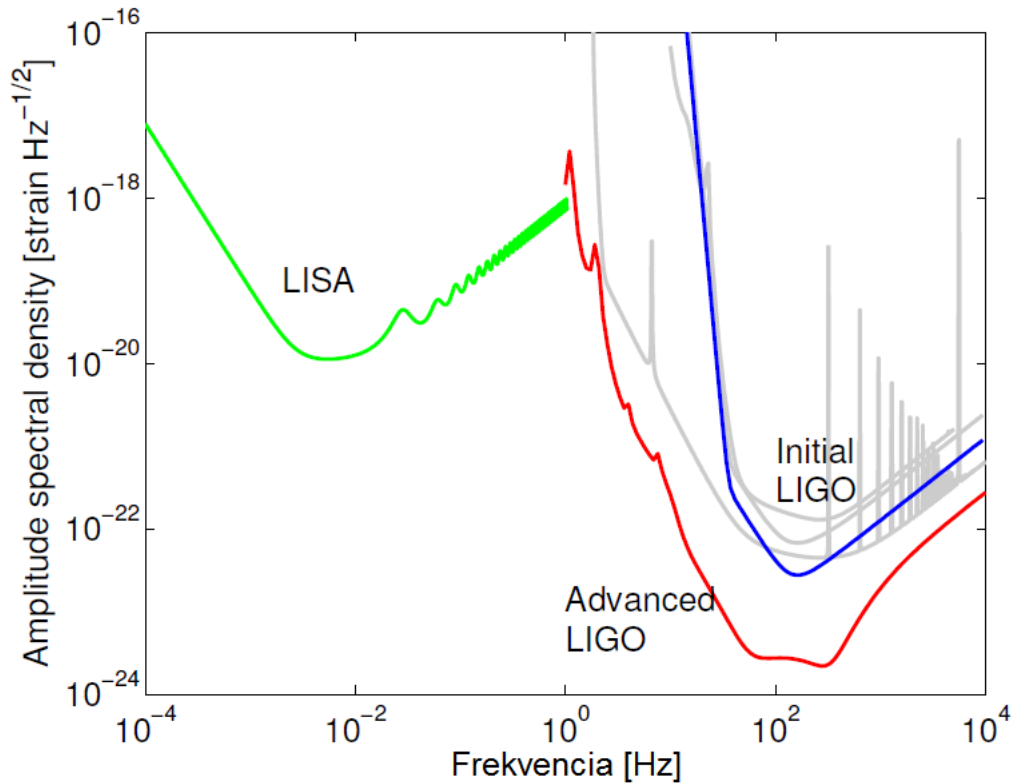
Egyes adatfeldolgozó programcsomagok (ún. koherens eljárások, lásd később) megkövetelhetik, hogy a digitális szűrés az adatpontok fázisát ne változtassa meg: ez megvalósítható azáltal, hogy az adatsornak a digitális szűrő időtérbeli reprezentációjával való konvolválása mellett a szűrő időtérbeli tükrözöttjével (azaz frekvenciatérbeli konjugáltjával) is megismételjük a műveletet. A kétszeres szűrés megőrzi a fázishelyességet, de a szűrésnek megfelelően



3.2. ábra. A LIGO hanfordi és livingstoni detektorainak zajspektrum sűrűsége, torzításban (*strain*) kifejezve. Az egyes görbék a különböző adatgyűjtési időszakok (*science run*) során elért legjobb eredményeket mutatják[15]. Szaggatott, illetve folytonos fekete vonallal jelöltük rendre a 2 és 4 km karhosszúságú detektorok tervezés során vett érzékenység-munkagörbéjét. Az alacsony frekvenciákon a szeizmikus zaj (földrétegek mozgásából fakad), míg a magasabbakon a sörétzaj (oka a fotonszám kvantumos fluktuációja) hatása meghatározó[9]. Az éles csúcsokat egyrészt a hálózati váltakozófeszültség, másrészt kalibrációs eljárások okozzák. (Forrás: [8]).

változtatja az adatsor amplitúdó spektrumát (fázistolásmentes szűrés, *Zero Phase Filtering* [23]).

A következő alfejezetben megvizsgáljuk az elterjedtebb adatfeldolgozó módszerek és programcsomagok főbb jellemzőit. Az algoritmusok közül csak azt tárgyaljuk részletesebben, amely közvetlenül kapcsolódik a munkánkhöz.



3.3. ábra. A jelenlegi és jövőbeni detektorok spektrális érzékenységgörbéi. A szürke görbesereget a 2 km-es LIGO, a GEO600 és a VIRGO detektorok tervezett érzékenységgörbéi alkotják (Forrás: [8]).

3.2. Keresőeljárások

3.2.1. Bázisválasztás

Általánosságban a keresőalgoritmusok működésében központi szerepet játszik az optimális bázis megválasztása: ugyanis legtöbbször a rögzített és előfeldolgozott adatokat egy megfelelően kiválasztott bázisra vetítik, majd a transzformált adatokban statisztikailag szignifikáns eltérést keresnek azon esethez képest, amikor a bemeneti jel nem tartalmaz GW-kitöréseket. A bázisválasztásnál figyelembe kell venni a keresett hullámforma sajátosságait. A választott bázis alapján megkülönböztethetünk a jelenlegi ismereteink alapján három jelkeresési eljárást³:

- **Párosított szűrés** (*Matched Filtering*) esetén a felismerni kívánt forrás jól modellezett, pontos hullámformák rendelkezésünkre állnak elméleti számítások alapján. Ezért az

³ Az egyes kategóriák között átfedések lehetségesek.

optimális bázis (vagy bázisok tere) ismert: az adatokat az ismert hullámformák alkotta bázison fejtjük ki.

- **Keresztkorreláció** (*Cross Correlation*, CC): több detektor által rögzített adatsorban jelen van ugyanazon, sokszor gyengén modellezett vagy teljesen ismeretlen hullámformájú forrás jele. Az egyik detektor adatait a másikéra vetítjük.
- **Absztrakt bázisok alkalmazása**: a várt hullámformák teljesen ismeretlenek. Ez esetben az adatokat a hullámformák egy megválasztott bázisára (ami lefedi a célzott jelek tartományát valamilyen paraméterterben) projektáljuk.

Ezen dolgozatban több detektor adatában, gyengén modellezett, vagy teljesen ismeretlen hullámformával jellemezhető *burst* típusú jelek keresésére koncentrálunk, ezért az általunk alkalmazott keresési eljárás központi eleme a keresztkorreláció.

Ez esetben az optimális bázis megtalálásának problémájára az algoritmus működésének ismertetésénél kitérünk.

3.2.2. Koherens és inkoherens módszerek

A fenti csoportosításon túl megkülönböztethetünk ún. koherens (*X Pipeline* [17], *Coherent Wave Burst* [18]) és inkoherens (*Omega Pipeline* [19]) adatfeldolgozó módszereket.

A **koherens** eljárások a jelkeresésben felhasználják a rögzített adatsor fázisspektrumát is, mivel több detektor adatsorát fázishelyesen (vö. koherencia) vetik össze. Az analízis során általában kettő gravitációshullám-detektor adatának idő-frekvencia térbeli reprezentációjának keresztkorrelációját (*cross correlation*⁴) határozzák meg. Az összegzésnél a fázishelyesség kritikus – egy valós asztrofizikai forrásból származó GW-jel fázisa eltérően projektálódna egymástól fizikailag távol telepített detektorok adatsorában. A koherens módszert csak *offline* (nem valós idejű) kereséseknél alkalmazzák magas erőforrásigénye miatt. Megjegyezzük, amennyiben valahonnan ismert a forrás égi pozíciója, ismert az időkéésés is, amely mellett az adatsorokat összehasonlíthatják – ezáltal a koherens keresés sokkal érzékenyebb lesz.

Az előbbiekkal szemben az **inkoherens** feldolgozás esetén az adatpontok fázisa nem hordoz releváns információt, mivel az algoritmus egyetlen detektor adatsorát elemzi. Tehát a jelkeresést csak az amplitúdó spektrum alapján végzik (a fázisspektrum ekkor természetesen irreleváns a keresés szempontjából). Az inkoherens analízis után lehetőség nyílik több detektor

⁴ Az elnevezésben szereplő *X* az alkalmazott keresztkorrelációra utal.

jelének összehasonlítására, koincidencia analízisre – így asztrofizikai forrás égi lokalizálására⁵ is[20].

A jelen dolgozatban az inkohereus algoritmusok analízisére, optimalizálására fókuszálunk. Ezt egyrészt gyorsaságuk, illetve a valós idejű kereséseknél, továbbá a pozíciórekonstrukcióban már betöltött – vagy további finomhangolás után potenciálisan betöltendő – meghatározó szerepük indokol. Másrészt az Omega Pipeline kódja kitörés típusú jelek keresésére optimalizált[19]. Az extragalaktikus asztrofizikai objektumok becsült száma és az Advanced LIGO detektorok tervezett érzékenysége alapján a jövőben várható nagyszámú kitörés detektálása. A következőkben az *Omega Pipeline* infrastruktúráját, továbbá a használt definíciókat vesszük górcső alá.

3.2.3. Az Omega Pipeline

Q-transzformáció

Az adatfeldolgozás elsődleges célja, hogy a bemeneti szűrt és simított spektrumú adatsor alapján – speciális bázis megválasztása után – előállítsunk kimenetként egy kétdimenziós idő-frekvencia diagramot⁶. Ezen képfeldolgozó algoritmusok alkalmazásával felismerhető a potenciálisan asztrofizikai forrásból származó GW-kitörések struktúrája⁷. A különböző detektorok spektrogramjainak összehasonlításával koincidencia analízis is végrehajtható (optimalizált algoritmusok esetén akár valós időben is).

A jelen analízis során a bemeneti adatok ún. idő-frekvencia-*Q* térbeli reprezentációját állítjuk elő, majd ezeket összefésüljük (*Q*-transzformáció).

A keresett kitörések paraméterezéséhez, az optimális bázis megtalálásához a következőkben bevezetünk egy reprezentációtól független, a gravitációshullám-kitörések nagyságrendjét jellemző mennyiséget (mértéket). Ez gyengén modellezett GW-kitörések mind keresése, mind detektálása esetén hasznosnak bizonyul, mivel lehetővé teszi néhány további segédmennyiség kézenfekvő definiálását. Ezen formalizmus alkalmazásával viszonylag kevés adattal jellemezhetünk idő-frekvencia térben jól lokalizált kitörés jellegű jeleket.

Ismert, hogy az idő és frekvencia térbeli leírás között a Fourier-transzformáció⁸ létesít

⁵ Az ún. pozíció rekonstrukció pontossága további detektorok alkalmazásával (az interferométerek, illetve a karok relatív pozíciójának megfelelő megválasztása mellett) jelentősen javítható.

⁶ Spektrogramot, azaz spektrumok sorozatát adott időablak esetén.

⁷ jellemzően időben gyorsan változó frekvencia, amelyre analitikus függvényalak nem ismert

⁸ A formulákban használt előjel konvenció a digitális jelfeldolgozás területén alkalmazottat követi.

kapcsolatot, tehát tetszőleges GW-kitörés esetén:

$$h(t) = \int_{-\infty}^{+\infty} \tilde{h}(f) e^{i2\pi ft} df \quad (3.3a)$$

$$\tilde{h}(f) = \int_{-\infty}^{+\infty} h(t) e^{-i2\pi ft} dt, \quad (3.3b)$$

ahol h a GW által okozott torzítást, f a frekvenciát, t az időt jelöli.

Továbbá tegyük fel, hogy a kitörések amplitúdója időben négyzetesen integrálható, ekkor definiálhatunk egy ún. karakterisztikus *strain* amplitúdót, $\|h\|$, amely kiszámítható egyaránt az idő és a frekvencia térben is⁹:

$$\|h\|^2 = \int_{-\infty}^{+\infty} |h(t)|^2 dt = \int_{-\infty}^{+\infty} |\tilde{h}(f)|^2 df. \quad (3.4)$$

Ezután egy tetszőleges GW-kitörés felírható az előbb bevezetett karakterisztikus *strain* amplitúdó és egy normált hullámfüggvény szorzataként:

$$h(t) = \|h\| \psi(t) \quad (3.5a)$$

$$\tilde{h}(f) = \|h\| \tilde{\psi}(f). \quad (3.5b)$$

Hasonlóan a nem-normált reprezentációkhoz, a normáltak is kifejezhetőek mind az idő, mind a frekvencia térben:

$$\psi(t) = \int_{-\infty}^{+\infty} \tilde{\psi}(f) e^{i2\pi ft} df \quad (3.6a)$$

$$\tilde{\psi}(f) = \int_{-\infty}^{+\infty} \psi(t) e^{-i2\pi ft} dt. \quad (3.6b)$$

A normált hullámfüggvények karakterisztikus strain amplitúdója definíció szerint egységnyi:

$$\int_{-\infty}^{+\infty} |\psi(t)|^2 dt = \int_{-\infty}^{+\infty} |\tilde{\psi}(f)|^2 df = 1. \quad (3.7)$$

A bevezetett normált hullámformák tehát amplitúdó-független leírását adják tetszőleges GW-kitörésnek.

A normált hullámfüggvények segítségével definiálhatjuk egy tetszőleges GW-kitörés τ

⁹ Az egyenlőség a Parseval-tétel miatt áll fenn[8].

centrális idejét és ϕ centrális frekvenciáját:

$$\tau = \int_{-\infty}^{+\infty} t |\psi(t)|^2 dt \quad (3.8a)$$

$$\phi = 2 \int_0^{+\infty} f |\tilde{\psi}(f)|^2 df. \quad (3.8b)$$

Az előbbi definíciókhoz hasonlóan bevezetjük egy tetszőleges GW-kitörés karakterisztikus időtartamát és sávszélességét az alábbiak szerint:

$$\sigma_t^2 = \int_{-\infty}^{+\infty} (t - \tau)^2 |\psi(t)|^2 dt \quad (3.9a)$$

$$\sigma_f^2 = 2 \int_0^{+\infty} (f - \phi)^2 |\tilde{\psi}(f)|^2 df. \quad (3.9b)$$

Végül definiálunk egy Q dimenziótlan ún. minőségi tényezőt, ami a centrális frekvencia és a sávszélesség aránya:

$$Q = \frac{\phi}{\sigma_f}. \quad (3.10)$$

Belátható, hogy a karakterisztikus időtartam és sávszélesség értéke között határozatlan-sági reláció áll fenn[21]:

$$\sigma_t \sigma_f \geq \frac{1}{4\pi}. \quad (3.11)$$

Azokat a kitöréseket, amelyekre a fenti egyenletben jó közelítéssel egyenlőség teljesül *jól lokalizált* kitöréseknek nevezzük.

Megjegyezzük, hogy a fent definiált idő-frekvencia térbeli paraméterek bár meghatározhatóak tetszőleges négyzetesen integrálható hullámfüggvény esetén, mégis legjobban jól lokalizált kitörések jellemzésére¹⁰ alkalmazhatóak[8].

Továbbá kézenfekvő bevezetni egy alternatív jelölést a kitörések idő-frekvencia térbeli időtartamára és sávszélességére:

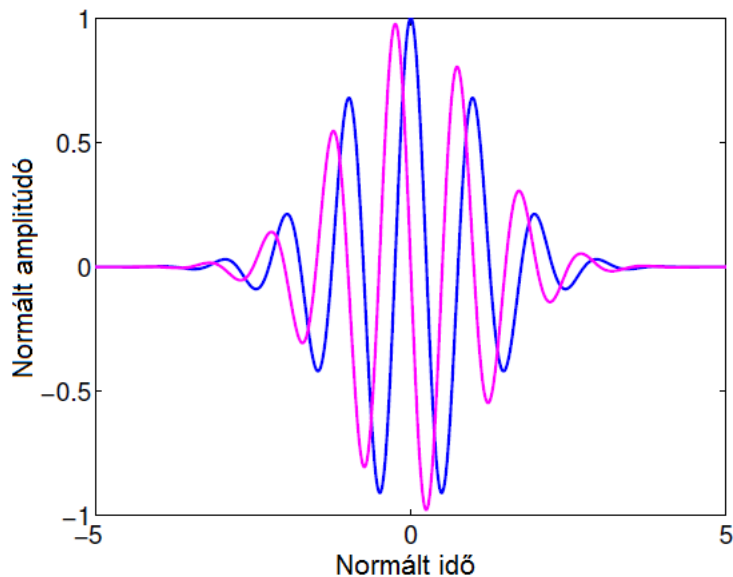
$$\Delta t = 2\pi^{1/2} \sigma_t \quad (3.12a)$$

$$\Delta f = 2\pi^{1/2} \sigma_f. \quad (3.12b)$$

Ezáltal a $\Delta t \Delta f$ szorzat egységnyi értéket vesz fel a minimális bizonytalanságú, jól lokalizált kitörés jelekre, a 3.11 egyenletnek megfelelően.

Továbbá belátható, hogy a 3.11 egyenlet által definiált téglalap területe akkor minimális, ha a vizsgált GW-kitörés idő- és frekvenciatérbeli reprezentációi ún. Gauss-modulált szinusz-

¹⁰Megjegyezzük, hogy az összetettebb burst-ök is kifejezhetők jól lokalizáltak lineáris kombinációjaként.



3.4. ábra. Normált Gauss-modulált szinusz függvény idő térben. A kék és magenta görbék a valós, illetve képzetes részét képezik a 3.13a és 3.13b egyenletekkel definiált komplex hullámformának.

függvény alakúak[21]:

$$h(t) = \|h\| \left(\frac{1}{2\pi\sigma_t^2} \right)^{1/4} e^{-\frac{(t-\tau)^2}{4\sigma_t^2}} e^{i2\pi\phi(t-\tau)} \quad (3.13a)$$

$$\tilde{h}(f) = \|h\| \left(\frac{1}{2\pi\sigma_f^2} \right)^{1/4} e^{-\frac{(f-\phi)^2}{4\sigma_f^2}} e^{i2\pi\tau(f-\phi)}. \quad (3.13b)$$

Ez a speciális jelalak nem-modellezett kitörések keresésénél kitüntetett jelentőséggel bír (köszönhetően a természetéből fakadó minimális bizonytalanságnak¹¹ és a viszonylag egyszerű, analitikus függvényalaknak).

A 3.4 ábrán láthatunk példát Gauss-modulált szinuszfüggvényre (1 Hz centrális frekvencia, 10-es minőségi faktor).

A Q -transzformáció során használt bázis megtalálásához átírjuk a 3.13a és 3.13b egyenletekkel definiált Gauss-modulált szinusz időtérbeli hullámformájú kitöréseket a normált alak-

¹¹ A Gauss-modulált szinuszfüggvény alakú hullámforma idő-frekvencia térben a legjobban lokalizált.

jukra (alkalmazva a fentebb definiált segédmenyiségeket):

$$\psi(t; \tau, \phi, Q) = \left(\frac{8\pi\phi^2}{Q^2} \right)^{1/4} e^{-\frac{4\pi^2\phi^2}{Q^2}(t-\tau)^2} e^{-i2\pi\phi(t-\tau)} \quad (3.14a)$$

$$\tilde{\psi}(t; \tau, \phi, Q) = \left(\frac{Q^2}{2\pi\phi^2} \right)^{1/4} e^{-\frac{Q^2}{4\phi^2}(f-\phi)^2} e^{-i2\pi\tau(f-\phi)}. \quad (3.14b)$$

Az átírásnál figyelembe vettük a 3.11 egyenletet, mely szerint az időtartam és a sávszélesség összefüggnek. Ennek eredményeképpen csak három paraméter szükséges a bázisfüggvényünk leírásához. Most a következőket választjuk: τ centrális idő, ϕ centrális frekvencia, Q minőségi faktor.

A folytonos Q -transzformáció a wavelet transzformáció általánosításának tekinthető: az $x(t)$ folytonos adatsor projekciója speciális komplex exponenciális függvények alkotta – ϕ centrális frekvenciával, Q minőségi faktorról jellemzett – teljes bázisra, folytonos esetben tehát:

$$X(\tau, \phi, Q) = \int_{-\infty}^{+\infty} x(t) w(t - \tau, \phi, Q) e^{-i2\pi\phi t} dt, \quad (3.15)$$

ahol $w(t - \tau, \phi, Q)$ egy időtérbeli ablakfüggvény (bázisfüggvény) τ időhöz centrálva, az időtartama arányos Q -val és fordítottan arányos a ϕ ún. centrális frekvenciával (jól lokalizált kitörések esetén).

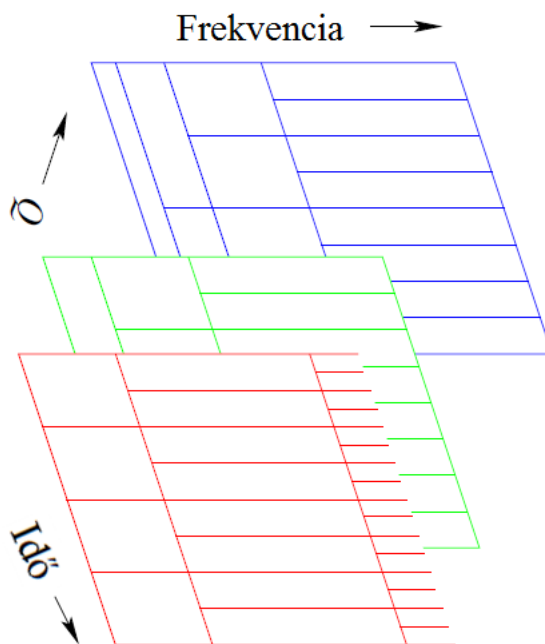
Elméleti megfontolások után adódik[8], hogy az optimális bázisfüggvényeknél az idő-frekvencia térbeli frekvenciaskálázás logaritmikus, az időtengely pedig lineáris. A bázisok Q paraméterei logaritmikusan követik egymást: ezáltal egy síksereget kapunk. Konstans Q faktor esetén a rácsozás hasonlít a wavelet transzformációnál alkalmazottra (3.5. ábra).

További adatkezelés

Több Q érték mellett is elvégezzük a fenti, wavelet dekompozícióra emlékeztető transzformációt, majd összefésüljük a különböző minőségi faktorhoz tartozó síkokat (**projekció** Q faktor szerint). Az eredmény egy különböző „energiájú” téglalapokat tartalmazó idő-frekvencia sík (térkép). Az adatsor Q -transzformációja után egy előre meghatározott energiaértéknél **vágást** hajtunk végre: elvetjük a túl alacsony „energiatartalmú” téglalapokat.

Az Omega Pipeline eljárás kimenete tehát téglalapok (ún. *event tiles*) idő-frekvencia térbeli pozícióit (t, f) és dimenzióit ($\Delta t, \Delta f$) definiáló adatokat tartalmazó vektorok serege.

Ezután a fenti folyamat eredményeképpen előálló téglalapokat a lokális sűrűségük alapján közös határoló téglalapokba foglaljuk – ezeket potenciális eseményeknek (*event candi-*



3.5. ábra. A Q -transzformáció speciális bázisának szemléltetése idő-frekvencia térképekkel. Konstans Q esetén az idő-frekvencia térbeli frekvenciaskálázás logaritmikus, az időtengely skálázása pedig lineáris. A különböző Q értékek logaritmikusán követik egymást.

date, EC; eseményjelölt) nevezzük. Az algoritmus kimenete jellemzően néhány darab EC koordinátája[19].

3.3. Koincidencia analízis

A fenti keresőeljárás eredménye detektoronként az EC-eket tartalmazó speciális spektrogram, az **eventgram** (potenciális eseményeket tartalmazó diagram; eseménytér(kép)).

A jelenleg implementált Omega Pipeline eljárás ezután az adott időintervallumokban előálló, kettő detektorból származó összes talált EC-t összefésüli. Amennyiben átfedést (koincidenciát) észlel a téglalapok között mind idő-, mind frekvenciaértékeket tekintve, akkor az algoritmus **riasztást** ad. Ez esetben további feldolgozáshoz (**utóanalízishez**) rögzítésre kerülnek a riasztást okozó EC-eket alkotó téglalapok centrális koordinátái és méretei az idő- és frekvenciatengely mentén.

Az automatikus feldolgozó eljárás következő lépése az ún. **Omega Scan**[25]. Ennek során történik egy adott időtartományban (általában 8 h alatt) rögzített három, legnagyobb energiátartalmú EC-eket tartalmazó riasztás (vagy esemény) detektoronkénti további elemzése. Ezeket a lokális fizikai környezetet monitorozó (*Physical Environment Monitoring*, PEM)

rendszer által előállított számos – a fentiekhez hasonló formátumú, különböző zajforrások járulékainak mintavételezésével generált – eseménytérképpel veti össze a pipeline soron következő programja. Ezen lépésben a fentebb részletezettel azonos koincidencia analízis valamely ismert zajforrás által triggerelt, azaz téves eseményeket hivatott kiszűrni. Az olyan eseményt, amely egyik detektor esetében sem tekinthető lokális zaj hatásának a program megjelöli (**kiemelt riasztás**). Az Omega Scan eredményeit a program valós időben feltölti egy weblapra[25]. Az adatokat jelszó megadása után az arra jogosultak megtekinthetik. A tudományos operátorok rendszeres jelentésekben foglalják össze az eredményeket.

4. fejezet

Képfeldolgozó eljárások

4.1. Motiváció

Az LIGO detektorokkal folytatott adatgyűjtés során a rendkívüli érzékenység mellett igen sok tényező játszik közre zajhatásként. Bár a zajszintet az évek során sikerült több nagyságrenddel csökkenteni, a GW-jelek direkt kimutatása továbbra is elérendő cél maradt. Az elméleti számítások alapján arra következtethetünk, hogy a GW-k anyaggal való rendkívül gyenge kölcsönható képessége miatt az első detektálható jelet a ma működő detektorok érzékenységének küszöbszintje környezetében kell keresnünk. Igen fontos cél, hogy az ilyen jelek adatsorbeli azonosításához minél érzékenyebb algoritmusokat dolgozzunk ki. Szoftveres adatfeldolgozó rendszerrel és jelkereső programokkal történő GW-jelek utáni kutatás önmagában is egyedülálló tudományos felfedezésre ad esélyt. A GW-jelek azonosítása konkrét asztrofizikai objektumok és folyamatok precízebb megértéséhez, modellek kidolgozásához vezetne. Ugyanakkor, tekintve az elméleti modellek sokaságát, a jelkeresési eljárások negatív eredménye már eddig is ezen modellek számát csökkentette, számos (köztük kozmológiai) modell szabad paramétereire pedig korlátokat szabott[22]. A fentiekből kitűnik, hogy érzékeny jelkereső algoritmusok kidolgozása, a már meglévők továbbfejlesztése számtalan téren és végeredmény esetén sikerrel kecsegtető célkitűzés.

A munkánk során a témaválasztást a fenti előnyök indokolták. Célunk egy olyan, az Omega Pipeline szoftver infrastruktúrára épülő program kidolgozása volt, amely alkalmazásával a keresőalgoritmus hatékonysága nagyban növelhető. A munkánkat az Omega Pipeline amerikai fejlesztőinek tudtával és támogatásával végeztük.

A LIGO detektorok adatsoraiban folytatott ún. *burst* keresés célja nem modellezett, viszonylag rövid időtartamú (< 1 s) GW-jelek azonosítása. Ez a keresés széles frekvenciatar-

ományban történik (80 – 2000 Hz), továbbá nem teszünk feltételezést a hullámformára és a forrás égi pozíciójára vonatkozóan sem. Ezért a korábban ismertetett matched filtering eljárás, továbbá egyéb, gyenge jel zajból való kiemelését célzó módszer alkalmazása nem célravezető. Az eddigi LIGO burst keresések inkohereus módszereken alapultak, amelyek során több detektor jelében kerestek egyezést az amplitúdó vagy az intenzitás (\sim energia) jellegű mennyiségek között (idő-, Fourier-, vagy wavelet-térben).

A rendelkezésünkre álló három LIGO detektor – amelyek irány-, polarizációérzékenysége és spektrális zajeloszlása hasonló – lehetővé teszik, hogy összehasonlítsuk a rögzített adatso-rokat és egyezéseket keressünk. Ezen célból – az Omega Pipeline keresőeljárást kiegészítendő – kifejlesztettük az **Overlapper** programot, amely alkalmazásával az eddigi, valódi jelke-resések során használt keresőalgorithmus hatékonysága nagymértékben javítható, illetve új keresési módszerek bevezetésével a gyengén modellezett GW-jelek detektálási esélyei nagy-ban növelhetők.

Az előző fejezetben ismertetett Omega Pipeline programcsomag kimenete jeltéglalapo-kat definiáló vektorok serege (egy detektorra, adott idő-frekvencia térben). Ismert, hogy az elemi eseményeknek egyértelműen megfeleltethetők téglalapok egy síkon (adott idő- és frek-venciatartományban), amelyek átfedhetnek egymással. A célunk különböző detektorok jó kö-zelítéssel azonos idő- és frekvenciatartományában rögzítésre került potenciális eseményeinek összehasonlítása (amelyeket a koincidencia ellenőrzés az algoritmusunk számára továbbadott és egymásnak megfeleltetett).

A vázolt ún. **koincidencia analízis** a tapasztalatok szerint meglehetősen sok téves ri-asztást produkál, azaz a detektorok adatsoraiban felismert koincidenciák az esetek döntő többségében a helyi zajhatások következményei. A téves riasztások nagy száma (magas ún. FAR, *False Alarm Rate*) miatt a koincidencia analízis a gyorsasága ellenére nem nevezhető optimálisnak.

A programunk kifejlesztéséhez e magas FAR érték redukálása szolgáltatta az elsődleges motivációt. Egy olyan jelkereső algoritmust dolgoztunk ki, amely EC-k összehasonlításán alapszik, és az idő- és frekvenciaértékbeli egybeesésen túl képes felismerni az esemény je-löltek alkotta komplexebb geometriai struktúrák hasonlóságát. Az eljárásból következően számottevően csökkenthető a véletlen koincidenciákból fakadó téves riasztások aránya, mivel a riasztás feltétele adott idő-frekvencia tartománybeli alakzatok kritikus **hasonlósága**. Ezen túl az algoritmusunk lehetővé teszi olyan események detektálását, amelyek valamilyen okból eltérő pozícióban¹ projektálódtak az idő-frekvencia síkra. Ez azért különösen fontos tényező a

¹ Az elméleti várakozások alapján az azonos forrásból származó jelek eltérő detektorok adatainak idő-

detektálás szempontjából, mivel a gravitációs hullámok véges terjedési sebességéből következően egy asztrofizikai forrásból származó kitörés jel a különböző interferométerek adatsorában adott, jól definiált időkéssel jelentkezne². A jelek beérkezési időpontjainak ismeretében a forrás éggömbi pozíciója bizonyos hibával meghatározható. A pozíciórekonstrukció hibája függ a rendelkezésre álló adatsorok, így a detektorok számától, illetve relatív karpozícióiktól.

A fentiek alapján megállapíthatjuk, hogy ismeretlen hullámformájú, nagyságrendileg másodperc karakterisztikus időtartamú (és időben gyorsan változó frekvenciájú³) események (4.1. ábra) keresésénél célravezető a GW-jelek adott síkbeli projekcióját alkotó jeltéglalapok⁴ egymáshoz mért hasonlóságát vizsgálni.

4.2. Az algoritmus

Az előző alfejezetben láttuk, hogy a GW-kitörések keresését visszavezettük egy képfeldolgozási probléma megoldására: kétdimenziós komplex alakzatok hasonlóságának mérésére. Megjegyezzük, hogy a jeltéglalapok energiaértékeit az analízis során nem használjuk fel, kizárólag összetett alakzatok hasonlóságát vizsgáljuk (tehát az összehasonlítás 2D-ben történik). Az algoritmusunk egy vétó eljárás: a megtalált EC-k közül eltávolítja azokat, amelyek nem hasonlítanak eléggé egymásra. A következőkben ismertetjük a jelenlegi algoritmus megtalálásához vezető fejlesztőmunka fontosabb állomásait.

4.2.1. Keresztkorreláció

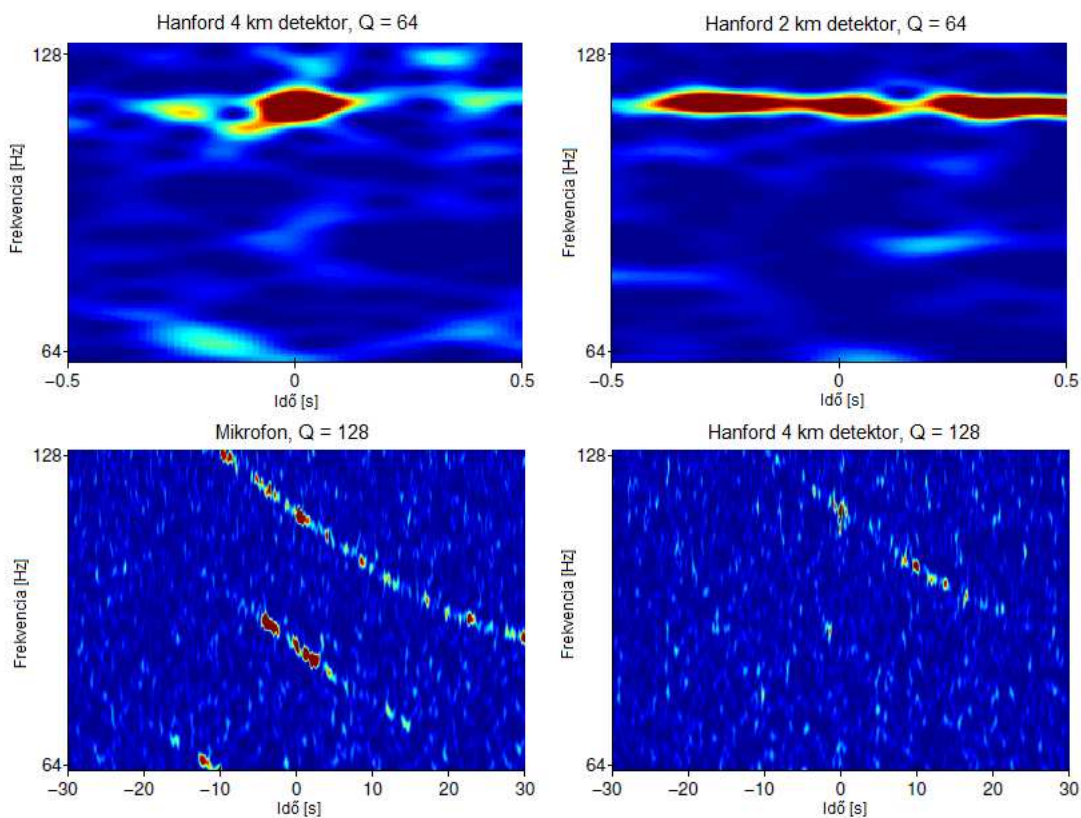
A Overlapper programunk inkoherens keresési eljárásról alapszik: azaz több detektor (esetünkben a 3 LIGO interferométer) adatsorainak szerves (koherens) egy sorba kombinálása helyett a kódunk keresztkorrelációt alkalmaz detektorpárok adatának összehasonlítására. További lépésekben újabb detektorpár(ok) adatsorát hasonlítja az előzőhöz. Ez a páronkénti összehasonlítás tetszőleges számú detektor adatsorára elvégezhető. Ezért a földi interferométerpárok teljes hálózatára elkészíthető a detektálási statisztika.

A fejlesztés során szem előtt tartottuk az algoritmus erőforrás-, illetve időigényének operatív frekvencia térbeli reprezentációjában hasonló alakú struktúrákat alkotnak.

² A LIGO detektorok közötti utat a fény – így várhatóan a GW is – 10 ms alatt teszi meg. Továbbá a karok eltérő pozícionálásából fakadóan – hacsak nem pontosan a detektorok síkjára merőlegesen érkezik be a GW – a detektorok irányérzékenysége eltérő.

³ Kvázi-monokromatikus állandó vagy lassan változó frekvenciájú gravitációs hullámok keresésénél jól alkalmazható a Locust- és a Hough-algoritmus[23])

⁴ Az idő-frekvencia sík viszonylag magas felbontása miatt várakozásaink szerint asztrofizikai források esetén a nagyszámú téglalapok kvázi egybefüggő alakzatot alkotnak.



4.1. ábra. Szignifikáns esemény konstans Q érték melletti idő-frekvencia spektrogramjai (Hanford, 2, illetve 4 km karhosszúságú detektorok). Az adatrögzítés során a detektor felett elszállt repülőgép okozta zajhatás eredménye a különböző karhosszúságú detektorok közötti koincidencia, ez az egyik monitorozó mikrofon adatsora alapján generált spektrogramon látható (lásd az alsó két ábrát). A frekvencia időbeli változása jól láthatóan a Doppler-effektusnak megfelelő (lásd az alsó két ábrát). A bal felső és jobb alsó ábrák ugyanazon spektrogramot mutatják, eltérő időtartamok és Q érték mellett. A felső és jobb alsó spektrogramok GW-csatorna adatából lettek előállítva, míg a bal alsó spektrogram a PEM egy mikrofonjának adatából. (Forrás: [8])

timalizálását. Törekedtünk olyan megvalósításra, amely valós idejű kereséseknél is alkalmazható (azaz online analízishez kellően gyors), viszont a téves riasztások aránya alacsony szinten tartható: így elektromágneses megfigyelések (follow-up) triggereléséhez is használható.

A LIGO Project legfontosabb rövidtávú célja GW-k közvetlen észlelése. Az egymástól 3002 km-re telepített detektorokból származó adatsorok kiértékelésénél keresztkorrelációs módszerek alkalmazása célravezető, mivel a statisztikailag korrelálatlan természetű, lokális eredetű zaj jó eséllyel eltávolítható, míg a korrelált, potenciálisan asztrofizikai forrásból származó jel kontrasztosodik, felerősíthető. Megjegyezzük, hogy a jelek korreláltságát nemcsak a közös, távoli forrás biztosítja, de a GW-k gyenge kölcsönhatása is az anyaggal, aminek következtében a jelek kvalitatív jellemzői (jelalak, jelhossz, sáv szélesség stb.) a detektorok között megtett távolság alatt nem változnak a rendelkezésünkre álló eszközökkel kimutatható mértékben.

Látható, hogy több, különböző detektor alkalmazása a GW-kutatásban meghatározó előnyökkel jár (például érzékenység növelése, forrás lokalizáció hibájának csökkentése). Fontos megjegyeznünk, hogy algoritmusunk alkalmas detektorok hálózataiból származó adatsorok gyors összehasonlító analízisére, így a jövőben telepítendő – mind érzékenyebb – földi detektorok által rögzített egyre növekvő adatmennyiség feldolgozására, szűrésére.

Az analízishez keresztkorrelációs módszert választottunk, mivel napjaink legérzékenyebb jelfeldolgozó eljárásainak számítanak. Az alkalmazásukkal jelkeresési szempontból igen rossz jel-zaj arány mellett is meglepően jó hatásokkal végezhetjük el a korrelált jelek kiemelését a zajos háttérből. Az alkalmazott módszer lényege, hogy különböző detektorok transzformált adatsorából két, azonos hosszúságú szegmens produktumával jellemezzük a korreláció mértékét. Esetünkben 2D-s idő-frekvencia diagramok hasonlóságának jellemzése a cél: ehhez először az előfeldolgozó analízis (Omega Pipeline) kimenetét alkotó eseményeket kell egy koordináta-rendszerbe leképeznünk, majd ez alapján a keresztkorrelációhoz mátrixot kell konstruálnunk detektoronként. A mátrixokat előállíthatjuk, ha az adatainkhoz egy rácsot illesztünk. Megjegyezzük, hogy ezen mátrixok csak 0 és 1 elemekből állnak. Ugyan az Omega Pipeline kimenete tartalmazza az események energiatartalmát is, az analízis során azt nem használjuk fel; az átfedő eseményektől eltekintünk. Ezután meghatározzuk a legalacsonyabb felbontású rácsot, amellyel mindkettő adat-sík alakzatai még tökéletesen felbonthatók. A felbontás elkészítése után előáll egy-egy mátrix, amely tartalmazza minden adatot az adatsíkok geometriájára vonatkozóan. Megjegyezzük, hogy ez esetben természetesen nincs információvesztés: az eredeti síkok diszkretizálása révén előállított mátrixok adattartalma azonos a

bemeneti fájlkéval, csupán egy másik reprezentációba transzformáltuk a bemeneti téglalap-seregeket.

Az imént tehát előállítottunk a bemeneti idő-frekvencia síkoknak megfelelő mátrixokat, úgy hogy ne veszítsünk információt. A következőkben a két mátrix hasonlóságát kívánjuk jellemezni egy számmal. Ehhez készítettük az **Overlap** nevű kódot, amely kiszámítja a mátrixok keresztkorrelációját: majd meghatározza a legnagyobb hasonlóság mértékét (CC) és hogy milyen relatív mátrix-pozíciók esetén sikerült azt elérni. Megjegyezzük, hogy az egydimenziós keresztkorreláció művelete folytonos f és g függvények esetén definíció szerint a következő integrállal fejezhető ki (a konvolúció speciális esetének tekinthető):

$$(f \star g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f^*(\tau) g(t + \tau) d\tau, \quad (4.1)$$

ahol f^* az f komplex konjugáltját jelöli. Diszkrét függvényekre, szintén egydimenziós esetre az előbbi definíció a következőképpen módosul:

$$(f \star g)[n] \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} f^*[m] g[n + m]. \quad (4.2)$$

Az általunk használt, kétdimenziós algoritmus a következő képletekkel adható meg:

$$C_{i,j} = \sum_{k=\max(1,i-c+1)}^{\min(a,i)} \sum_{l=\max(1,j-d+1)}^{\min(b,j)} A_{k,l} \cdot B_{c-i+k,d-j+l} \quad (4.3)$$

$$CC = \max(C_{i,j}), \quad (4.4)$$

ahol $C_{i,j}$ az ún. korrelációs mátrix, A és B az összehasonlítandó mátrixok, a és b rendre az A mátrix sorainak és oszlopainak száma, c és d rendre az B mátrix sorainak és oszlopainak száma, CC a keresett normálatlan keresztkorreláció.

Végül az ún. normált keresztkorrelációt (normált CC , NCC) állapítjuk meg – az egyes hasonlóság értékek összevethetősége érdekében, ami a következő alakban definiálható:

$$NCC = \frac{CC}{\|X\| \cdot \|Y\|}, \quad 0 \leq NCC \leq 1, \quad (4.5)$$

ahol az X és Y az egyes bemeneti mátrixok elemeit jelölik, a $\|\cdot\|$ pedig az L^2 norma. Megjegyezzük, hogy a normálási faktorok a mátrixok reprezentálta alakzatok önmagukkal vett hasonlóságának abszolút mértékével (CC) egyenlők.

A detektorok rendkívüli érzékenysége miatt az előálló idő-frekvencia diagramok igen nagy felbontásúak: leképezésük rácsra nem optimális, mivel a keletkező mátrixok dimenziója nagyságrendileg a 10^5 értéket is elérheti, elemszámuk megközelíti a több milliárdot, így azok numerikusan kezelhetetlenek.

4.2.2. Az iterációs eljárás

Tehát az Omega Pipeline folytonosan megadott téglalapjaiból egy diszkrét rácsozással egy mátrixot állítunk elő, amit 0 és 1 számértékek alkotnak. Ez nem működik a téglalapokhoz tökéletesen illeszkedő ráccsal, mert kezelhetetlenül nagyok lesznek a mátrixok. Ehelyett nem tökéletesen illeszkedő rácsozást használunk és próbáljuk közelíteni az eredményünket az egzakt rácsozásból kapható eredményhez. Olyan módszert kell tehát választani, ami határ esetben visszaadja a tökéletes (natív) rácsozás végeredményét.

Amennyiben felhagyunk az egzakt számítás megkövetelésével, a probléma numerikusan kezelhetővé válik – a rendelkezésünkre álló számítási erőforrásokkal is – sőt, igen jelentős sebességnövekedés érhető el a futási időt tekintve. Az optimalizálás kézenfekvő útja iterációs algoritmus alkalmazása.

Az iterációs módszer alkalmazásának vitathatatlan előnye a skálázható futási idő (az iterációs szám megválasztásával), illetve a kontrollálható erőforrásigény (RAM, háttértár). Olyan iterációs eljárás kifejlesztését céloztuk meg, amelynél az összehasonlítandó mátrixok maximális mérete megadható, ezáltal lehetőség nyílik a program valós idejű kereséseknél való használatára. Az iterációs szám helyes megválasztásával a futási idő – információveszteség⁵ fordítottan arányos mennyiségeket kell egyszerre minimalizálnunk; azaz egy ésszerű kompromisszumot kötni. A következőkben ismertetjük a kifejlesztett iterációs algoritmust.

Az iterációs algoritmus szintén a kétdimenziós keresztkorrelációs eljárásról alapszik (lásd Overlap), viszont megegyező méretű (négyzetes) mátrixokat hasonlítunk össze.

Egy detektorpár esetén nem feltétlenül egyezik meg a teljes idő-frekvencia síkok területe, ezért az analízisünk első lépéseként a két detektor adatfájljait összefésüljük, és meghatározzuk azon legkisebb közös határoló téglalap koordinátáit, amely bármely eseményhez rendelt téglalapot tartalmazza. Ezzel megadjuk az adott analízis értelmezési tartományát. Továbbá elértük, hogy a korrelációs analízis bemeneti mátrixai azonos felbontásúak legyenek, amely

⁵ Az információveszteség értelmezhető extra statisztikus zajforrásként a feldolgozás során; a mintaszám – így a futási idő – növelésével csökkenthető.

az összehasonlíthatóság feltétele. Ez azt eredményezi, hogy az egyes mátrixelemek mindkét detektor esetén ugyanakkora területű idő-frekvencia síkbeli szegmensnek kell, hogy megfelelően legyenek: röviden, szükséges, hogy a mátrixok ún. *pixel*méretei megegyezzenek.

Ezután elhelyezzük mindkét téglalap sorozatot az imént meghatározott határoló (közös) legkisebb téglalapba, majd külön-külön mindkettő – immár megegyező méretű – téglalapot felosszuk $2^i \times 2^i$ pixelre minden iterációs lépésben, ahol i az iterációk számát jelöli. Egy adott felosztásnál előállítunk egy-egy $2^i \times 2^i$ méretű mátrixot: meghatározzuk egy éppen aktuálisan kiszemelt pixellel átfedő téglalapterületek összegének⁶ arányát a pixel teljes területéhez képest. Az így, minden iterációs lépésben előállított mátrixpárt keresztkorrelációnak⁷ vetjük alá (lásd a 4.2.1. alfejezetet), és a normált CC értéket eltároljuk.

A fentebb vázolt lépéseket előre definiált N iterációszámig elvégezzük, majd a normált CC értékek sorozatára a következő alakú exponenciális függvényt illesztjük:

$$f(x) = A + B \cdot e^{-\lambda x}, \quad (4.6)$$

ahol az x független változó az iterációszámot jelöli; A , B , λ pedig illesztési paraméterek. Azt az A értéket, amelyhez az illesztett görbe konvergál, tekintjük a komplex alakzatok – extrapolált – hasonlósági mértékének.

Megjegyezzük, hogy a logaritmikus-lineáris skálán történő egyenesillesztés esetén a fenti A konstans paraméter meghatározása nehézkes. Természetesen az alkalmazott nemlineáris illesztés is rendelkezik hátrányokkal: különösen problémás lehet a fenti illesztési paraméterek kezdeti értékének megfelelő beállítása. Előfordulhat ugyanis (főleg akkor, ha nagy a statisztikus hibája az adatpontoknak), hogy a nemlineáris illesztés nem vagy helytelenül konvergál. Az exponenciális illesztés javítása céljából saját illesztőalgoritmust fejlesztettünk ki.

Továbbá megjegyezzük, hogy a különböző tesztheink során gyakran tapasztaltunk olyan adatsorokat, amelyre az illesztés nem, vagy csak nehézségek árán volt megvalósítható – ez azonban nem vonja maga után, hogy az általunk feltételezett modell hiányos, vagy hibás lenne. Sőt, mivel a nem konvergáló illesztések szinte kivétel nélkül erősen különböző alakzatoknál következtek be, így őket egyfajta indikátornak tekinthetjük. Az általunk használt interaktív MATLAB fejlesztési környezetben beállítható a maximális függvénykiértékelések

⁶ A többszörös átfedések ignorálhatók, mivel valós keresésnél az általuk okozott extra járulék elhanyagolható, továbbá belátható, hogy a detektálási statisztikát szignifikáns mértékben nem módosítanak.

⁷ A keresztkorreláció művelete a $2-2^7$ méretű mátrixokra összesen szinte minden vizsgált esetben kevesebb, mint 10 s-ig tart.

száma: így az elméleti modellel erősen inkompatibilis adatsorok nem idéznek elő adattorlódást a GW-kereső pipeline-ban, hanem elvetésre kerülnek. Így bizonyos rosszul definiált klaszterelrendezések hatására nem kerül végtelen ciklusba az adatfeldolgozó eljárás: legrosszabb esetben egy perc (átlagosan 30 s) alatt elvetésre kerül a fennakadást előidéző spektrogram.

A paraméter-inicializációval kapcsolatban megjegyezzük, hogy a jelenlegi algoritmusunk képes klaszter-elrendezések igen széles tartományára automatikusan helyes kezdeti értéket adni az illesztési paramétereknek.

A fentieknek megfelelően az első iterációk kiszámítása gyors, ámde azok jelentős statisztikus hibával terhelt⁸ becslései az egzakt normált hasonlósági mértéknek. A végzett tesztjeink alapján az algoritmus időfüggése az iterációs mélységgel magas (3, 4) kitevőjű polinomfüggvény szerint skálázik. Tehát az egymást követő iterációk időigénye ugrásszerűen növekszik.

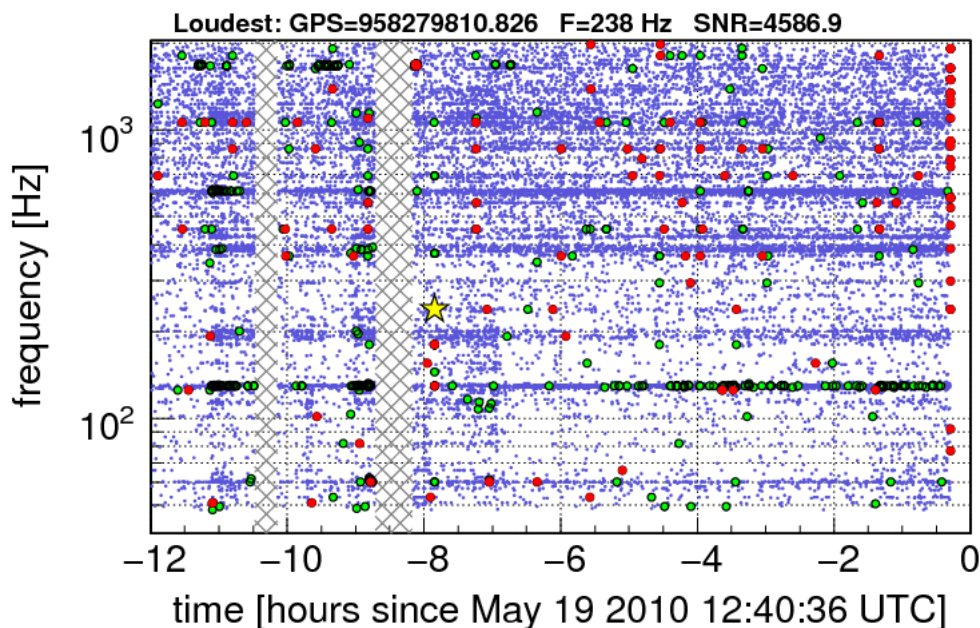
Tesztjeink alapján a legtöbb esetben – a bemeneti téglalapok száma nagyságrendileg 10 és 1000 között volt – az 8 iteráció tűnt optimális választásnak (futási idő < 1 min a rendelkezésre álló számítási kapacitás mellett).

Felmerült a kérdés, mennyiben szükséges a pontosság megkövetelése a paraméterbecslésnél, hiszen algoritmusunk egy vétó eljárás szerepét tölti be; a kérdés alapos tisztázáshoz több típusú tesztnek vetettük alá az eljárásunkat.

Az algoritmusunk a normált CC érték mellett képes meghatározni a „legtökéletesebb” átfedéshez (*overlap*; amikor a hasonlóság maximális) tartozó relatív mátrixpozíciót (*Shift*). Ezek meghatározása szintén iterációs eljárásról alapszik. Viszont exponenciális illesztést csak a hasonlósági mértékre vonatkozóan végzünk: a *Shift* értékek konvergenciája ugyanis a konkrét alakzatoktól függ. Ezért a programunk jelen verziója a *Shift* változó utolsó iteráció során felvett értékét tekinti a legjobb becslésnek.

Megjegyezzük, hogy amennyiben a keresztkorrelációs analízist csak idő dimenzióban végezzük el – ha eltekinthetünk a frekvencia változásától –, akkor jelentősen csökken a futási idő, így pontosabb hasonlóságbecslés érhető el. További optimalizálási lehetőség lehet, ha az egyes iterációknál figyelembe vesszük az előzőben kapott maximális hasonlóság értéket, és az

⁸ Megjegyezzük, hogy kísérleteztünk súlyozott nemlineáris illesztést megvalósító algoritmusokkal is – elméleti modellek alapján növekvő súllyal vettük figyelembe az egymást követő iterációk eredményét. Ám egyik sem kecsegtetett potenciálisan annyira jelentős előnnyel, hogy a végleges verzióban implementálásra kerüljön. Továbbá utalnánk rá, hogy az algoritmusunk egy vétó eljárás szerepét hivatott betölteni és a valós detektálásnál elsősorban többségben vannak az elégtelen hasonlóságú, vagy illesztési modellel inkompatibilis adatsorok.



4.2. ábra. 2010 májusában az Omega Pipeline keresőeljárás által rögzített eventgram. Minden színes pont egy, az Omega Pipeline által megtalált esemény. A színek jelzik, hogy az esemény jel-zaj aránya (SNR) milyen tartományba esik. Kék: $5 < SNR < 10$, zöld: $10 < SNR < 20$, piros: $SNR > 20$. A csillag a 12 h alatt rögzített legmagasabb jel-zaj arányt mutatja. (Forrás: [25])

ahhoz tartozó relatív időeltolásnak (ha eltekintünk a frekvenciabeli változástól) csak bizonyos időkörnyezetében folytatjuk az összehasonlítást.

A fentiek alapján megállapíthatjuk, hogy az iterációs algoritmusunk valós idejű keresésnél (4.2. ábra) hatékonyan alkalmazható, ezt a következő tulajdonságai indokolják:

- Gyors futás, a keresés szempontjából pontos hasonlóságbecslés;
- Skálázható pontosság a rendelkezésre álló erőforrásoknak megfelelően;
- Optimalizált nemlineáris illesztés, hatékony paraméterbecslés.

4.3. Tesztek

Az előző alfejezetben utaltunk rá, hogy számos tesztfuttatást végeztünk bizonyítandó az eddig implementált algoritmusunk hatékonyságát. Itt jegyezzük meg, hogy a fejlesztés során külföldi kutatókkal szorosán együttműködünk.

Továbbá fontosnak tartjuk megjegyezni, hogy a következőkben ismertetésre kerülő tesztelő eljárás saját ötleten alapszik, és tudomásunk szerint még nem implementálták.

Amennyiben külön nem jelezzük, a következő tesztek kliens PC-n (Intel Celeron Mobile egymagos processzor, 1 GB RAM) futtattuk.

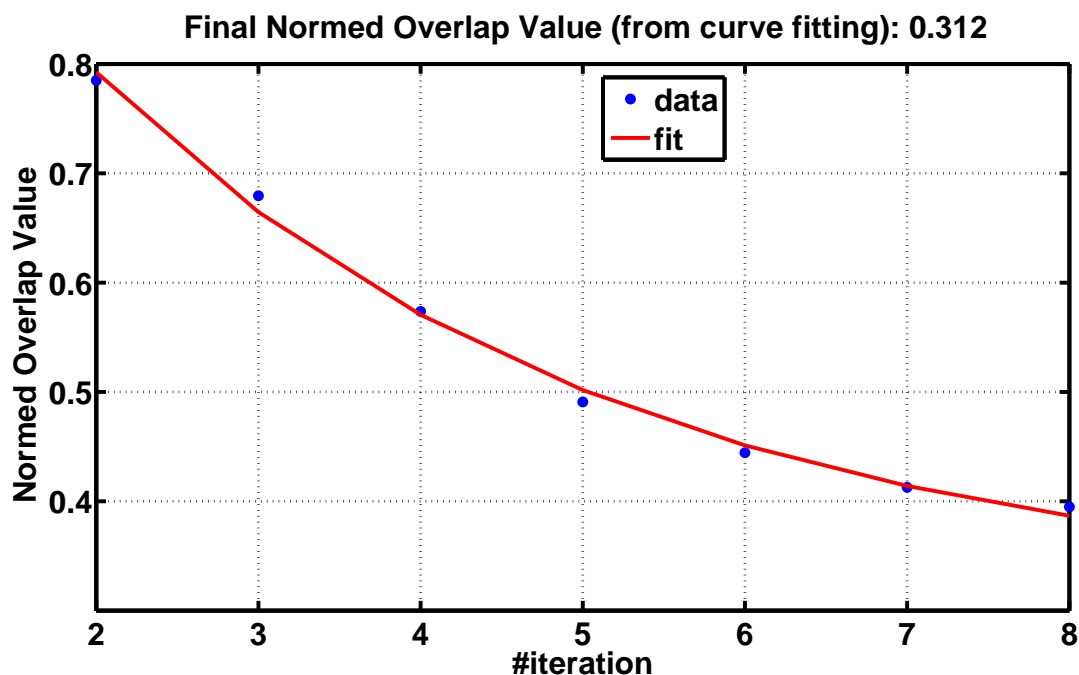
Volt alkalmunk kipróbálni a programunkat valódi, Omega Pipeline kimeneti adatokon is. A tapasztalataink azt mutatják, hogy a hasonlóság korrekt becsléséhez $\approx 8 - 10$ iteráció szükséges. A további tesztekben is ezt az optimálisnak tekinthető iterációszámot használtuk, mivel ekkor az általunk vizsgált esetekben szinte kivétel nélkül rövidebb volt a futási idő egy percnél, ugyanakkor a tesztek döntő többségében a programunk felismerte az egybevágó, illetve hasonló alakzatokat.

Az algoritmusunk tesztelése érdekében kidolgoztunk további programokat. Ezekkel lehetővé vált, hogy az Omega Pipeline kimeneti fájlformátumával megegyező struktúrájú adatsorokat generáljunk bitképek alapján. Az Omega Pipeline infrastruktúrájának megfelelően az általunk készített bitképek fekete-fehérek és 200×100 pixel felbontásúak voltak. Ezen veszteségmenetes BMP képeket átkonvertáltuk PBM formátumúra. Ezen ASCII formátumú fájlokból előállítottunk Omega Pipeline konform adatsorokat: konvenció szerint az '1' értékű pixeleket tekintettük potenciális eseménynek. Egy program megkereste az adott értékű pixeleket, majd azok koordinátáit egy szöveges fájlban rögzítette. A valódi Omega formátumtól az eljárásunk eredményeként előálló fájl annyiban tért el, hogy esetünkben a pixelek négyzetesek, mindkét oldaluk egységnyi; továbbá a pixelek egymással nem fedhetnek át. Megjegyezzük, hogy ezen eltérés nem jelentős, ugyanis a valódi Omega fájlokban a téglalapok átfedése nem gyakori.

A tesztekhez készített első bitképek egyszerű geometriai alakzatokat (téglalap, kör) ábrázoltak; majd JPEG fotók alapján készítettünk összetett alakzatokat (arckép) ábrázoló bitképeket is.

A tesztek során elvégeztük:

1. Egybevágó alakzatok (például eltérő pozíciójú téglalapok) (4.5. ábra);



4.3. ábra. Portrék (lásd a 4.6. ábrát) normált hasonlóságának alakulása az iterációs szám függvényében. Az illesztéssel meghatározott normált CC értéknek 0.312 adódott.

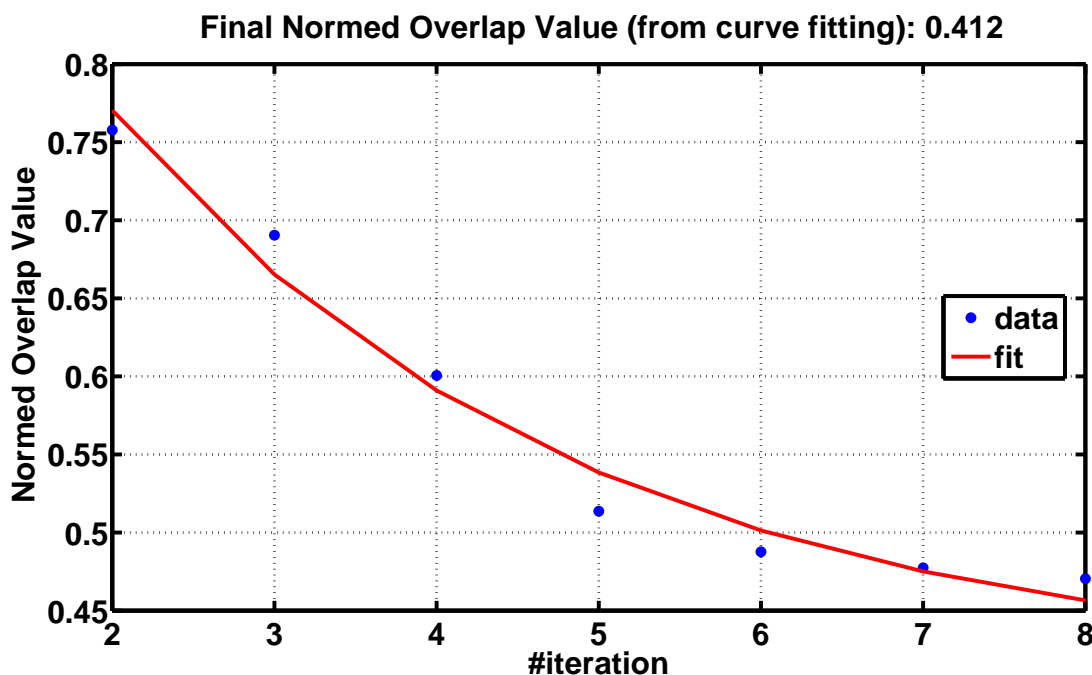
2. Eltérő alakzatok (például kör és téglalap) (4.5. ábra);
3. Átfedő alakzatok (például kör, illetve átlapoló kör és téglalap) (4.5. ábra).
4. Nagyobb méretű, komplex, nem egybefüggő bitképek (például portrék; lásd a 4.3., 4.4. és 4.6. ábrákat)

összehasonlító elemzését.

Megjegyezzük, hogy ezen összehasonlítások során a programunk egyszerre átlagosan 4–8 ezer téglalapot dolgoz fel bemeneti fájlként, szinte minden esetben 30 s alatt. Ekkor az utolsó iterációs lépésben már 65536 elemű mátrixok összehasonlítása történik.

A normált CC értékre a következő eredményeket kaptuk:

1. Tökéletes egyezésnél: $CC_{eq} = 0.989$
2. Eltérő alakzatoknál: $CC_{dif} = 0.540$
3. Részben átfedő alakzatoknál: $CC_{ol} = 0.806$
4. Portrék összehasonlításánál: $CC_{comp} = 0.312$



4.4. ábra. Alakzatfelismerés során a normált hasonlóság alakulása az iterációszám függvényében. Egy arc részlet keresését végezzük zajos, több portrét részben összefésülve tartalmazó képen (4.6. ábra). Az illesztéssel meghatározott normált CC értéknek 0.412 adódott.

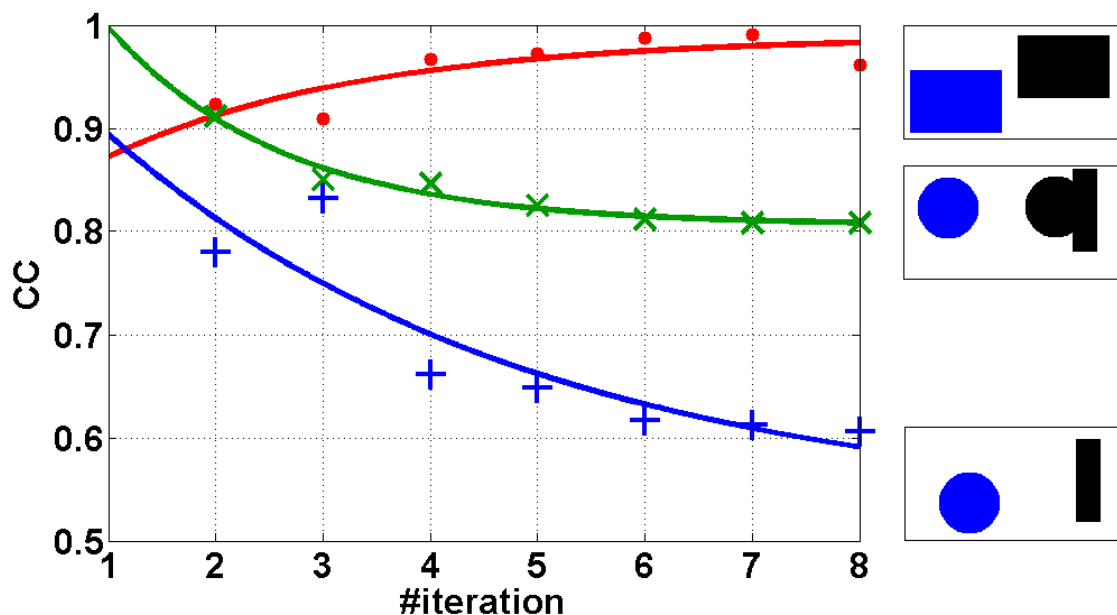
A portrék összehasonlítása rávilágít, hogy igen komplex alakzatok hasonlóságának jellemzésére egy szám könnyen elégtelen lehet.

A tesztek alapján levonhatunk néhány fontos konklúziót:

1. Megegyező alakzatok esetén a jól illeszkedő exponenciális függvénygörbe igen jól közelítette az 1 hasonlóság értéket.
2. Eltérő alakzatok esetén a rosszul, vagy egyáltalán nem illeszkedő görbe viszonylag gyorsan lecsengett – vagy az illesztés eleve nem konvergált.
3. Részleges egyezés esetén a pontsorra jól illeszkedő görbe szintén lecsengett, ám viszonylag magas hasonlóság értékhez ($NCC \approx 0.80$) konvergált.

A fenti teszteredmények alapján az illesztési algoritmusunkat tovább finomhangoltuk: a programunk megvizsgálja, hogy monoton növekvő, vagy csökkenő a pontsor lefutása (az első és utolsó iterációban számolt pont értékének összehasonlításával):

- amennyiben növekvő, akkor a korábban ismertetett illesztést mindenképpen végrehajtja (jó eséllyel igen hasonló alakzatokra számítunk);



4.5. ábra. A tesztjeink eredményei összesítve: különböző klaszterelrendezések esetén a normált hasonlóság alakulása az iterációs szám függvényében. Az egyes illesztett görbékhez a jobb oldali klaszterelrendezések tartoznak. Az extrapolált normált hasonlóság értékre az azonos, részben átfedő, illetve különböző alakzatoknál rendre 0.989, 0.806, 0.540 adódik.

- csökkenő lefutás esetén,
 - ha az utolsó pont értéke egy előre definiált p paraméter alatti, akkor a hasonlósági mértéknek definíció szerint az utolsó pont értékét tekinti (gyenge hasonlóság, potenciálisan rossz illeszkedés az elméleti modellhez);
 - amennyiben a p paraméternél nagyobb, akkor az illesztést végrehajtja (átfedő alakzatokra számítunk).

A p paraméter értéke a program futtatásakor megadható, mi a $p = 0.50$, 0.60 értékeket alkalmaztuk.

Továbbá a programunk minden esetben megtalálta a „legjobb” átfedését a vizsgált képeknek (amikor a hasonlóság maximális), az arcképek esetében is. A portrék összevetése során megtalált legjobb illeszkedést összevetve a két kép közvetlen összefésülésével látható, hogy portrék azonos részei ténylegesen precízebben illeszkednek, amennyiben a programunk kimenetének megfelelő relatív pozícióban fésüljük őket össze.

A tesztek során felismertük, hogy a képek keresztkorrelációja során meghatározott „legjobb illeszkedés” koordinátái (*Shift*) alapján az algoritmusunk alkalmas alakzatfelismerésre



4.6. ábra. Az alakzatfelismerés teszteléséhez használt képek (fent, jobbra a keresett arc-részlet, balra a zajos háttérkép), és a futtatás eredményeként kapott legjobb illeszkedéshez tartozó pozíció (lent).

is. Ezt egymással átlapoló portrékban (kvázi zajos háttérben) egy arc-részlet megkeresésével szemléltetjük: az általunk kiszemelt – kis mértékben torzított – mintát tökéletesen megtalálta az algoritmusunk a zajos háttérképben. A sikeres és gyors alakzatfelismerés újabb alkalmazási területeket nyit meg az algoritmusunk előtt, többek között a gravitációshullám-asztrófizika területén is.

Az **alakzatfelismerés** lehetővé tette egy új keresési módszert is: ezáltal algoritmusunk ismeretlen hullámformájú burst-ök keresésén túl alkalmas ismert, viszonylag jól modellezett jelalakok zajos jelben történő gyors kimutatására.

Ezáltal a jövőben lehetővé válna – megnövekedett detektorérzékenység mellett – bizonyos típusú (például nemzérus spinű) bespirálózó kettősök jelének szűrése a rögzített adatsorból. Az általunk javasolt eljárás lényegében matched filtering végrehajtása egy speciális wavelet térben.

Láttuk, hogy mintafelismerésre a programunk zajos háttér, illetve hiányos modell esetén is képes. Ez azért különösen lényeges szempont, mivel az elméleti becslések alapján a GW-kitöréseket a jelenlegi detektorok érzékenységi küszöbszintjének kis környezetében érdemes

keresnünk. Tehát jó eséllyel a rögzített jel nem egybefüggő alakzatként projektálódik a keresési térbe, hanem „széttöredezik”. Ezért kitüntetett jelentősége van jó hibátűrű, redundáns keresőalgoritmusok kifejlesztésének.

Továbbá az alakzat megtalálásának esélyei javíthatók:

- további detektorok alkalmazásával;
- modellek egy tartományának párhuzamos keresésével;
- további alakzatfelismerő módszerek (például a Locust-algoritmus[23]) alkalmazásával – amellyel különösen wavelet térben polinomfüggvénnyel jól modellezhető lefutású jelek alakja rekonstruálható, így a zajos háttértől kontrasztosabbá váló jel a programunk által könnyen kiemelhető.

Továbbá érdemes megjegyezni, hogy jelenleg az Omega Pipeline infrastruktúrában alkalmazott lokális sűrűség alapú klaszterező eljárás nem optimális. A kutatócsoportunk jelenleg is egy hatékonyabb klaszterező algoritmus kifejlesztésén dolgozik; illetve vizsgáljuk e téren más képfelismerő algoritmusok párhuzamos alkalmazásának hatékonyságát.

4.4. Szimuláció

A programunkkal végeztünk tesztek szimulált bemeneti fájlok használatával is. A szimulált zajjal végzett tesztjeinket az algoritmusunk – különböző esetekre érvényes – **téves riasztási** arányának (elsőfajú statisztikai hiba: nincs jel, mégis riaszt a program), az ún. FAR (*False Alarm Rate*) érték meghatározása motiválta. A zaj, mint statisztikai folyamat eloszlásfüggvénnyel jellemezhető: esetünkben az egyenletes eloszlást választottuk a zajhatást reprezentáló téglalapok pozícióinak (t centrális idő, f centrális frekvencia) és méretének (σ_t karakterisztikus időtartam, σ_f karakterisztikus sávszélesség) álvéletlen értékeinek generálásához.

Ezáltal tehát véletlenszerűen előállított eseménysorozatokat – azaz szimulált zajt⁹ – tartalmazó eseménytérképeket hasonlítottunk össze. Az Omega Pipeline kompatibilis adatsorokat előállító eljárást a következőkben ismertetjük.

Az ezen teszthez kifejlesztett programunk nagyszámú Omega formátumú adathalmazt generált, majd azokat páronként az **Overlapper**-kód összehasonlította, majd a hasonlóság mértékét, illetve a legjobb hasonlósági mértékhez tartozó relatív pozícióadatokat egy fájlban rögzítette.

A szimulációs környezet bemeneti paraméterei a következők:

- N , a generálandó adathalmazok száma;
- az ún. η faktor, amellyel egyenesen arányos az egyes futásonként összehasonlítandó spektrogramok eseményeinek száma¹⁰.

A teszteléshez használt spektrogramokat négyzetesnek választottuk, a oldalhosszal¹¹.

Megjegyezzük, hogy az Omega Pipeline infrastruktúrájában alkalmazott Q -transzformáció által meghatározott bázis sajátosságai miatt a frekvenciaértékeket logaritmikusan skálán, egyenletes eloszlással generáltuk, majd azokat exponencializáltuk.

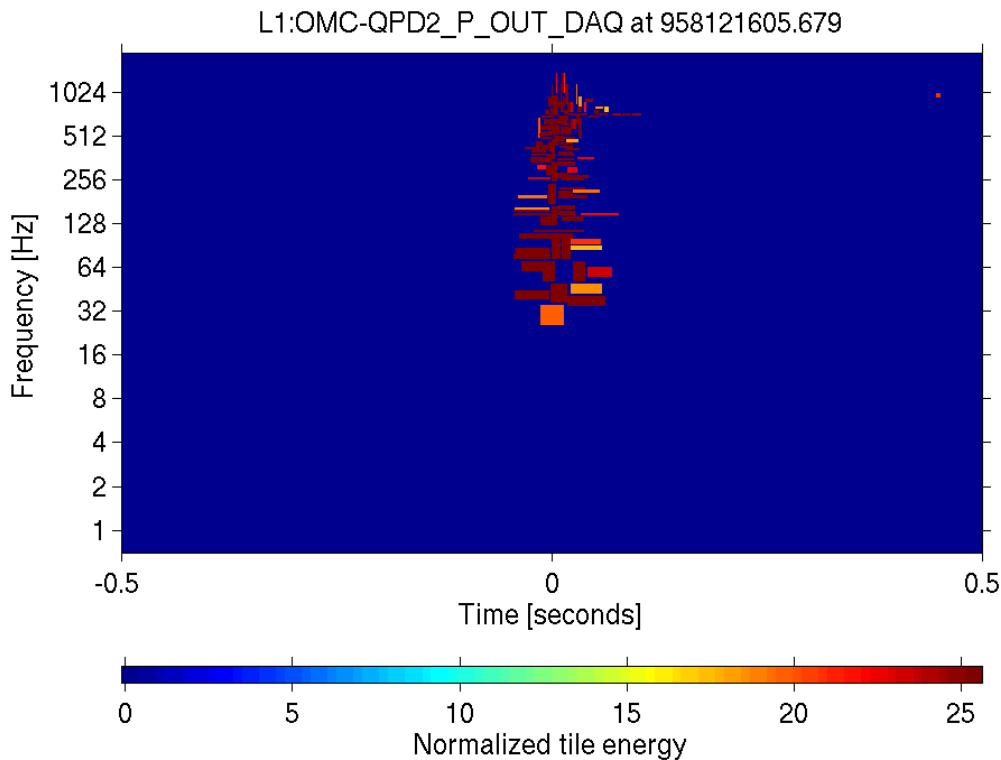
A tesztjeinkhez jól lokalizált eseményeket tartalmazó spektrogramokat generáltunk, azaz a 3.11 egyenletnek megfelelően az idő- és frekvenciabizonytalanság szorzata (a téglalapok területe) minden esetben $(4\pi)^{-1}$ egységnyi volt, tehát a jelöléseinkkel:

$$\Delta t \Delta f \stackrel{!}{=} \frac{1}{4\pi}. \quad (4.7)$$

⁹ Az egyenletes frekvenciaeloszlás miatt fehér zajnak tekinthetjük.

¹⁰ Az eseményszám megegyezik a generálandó Omega fájlparók sorainak számával.

¹¹ Az a értékét elméleti megfontolások után 10 egységnek választottuk.



4.7. ábra. A 4 km karhosszúságú livingstoni detektor adatsorán valósidőben futtatott Omega Pipeline eljárás által előállított eventgram másodperces szelete. Az algoritmusunk ilyen téglalapsorozatotak hasonlít össze. (Forrás: [25])

Ezáltal a tesztjeink értelmezési tartományaként szolgáló spektrogramok téglalapokkal való lefedettségi aránya meghatározható, $4\pi a^2$ -tel egyenlő. A lefedettségi százalékkal arányos természetesen a téglalapok száma (az átfedő téglalapoktól most eltekintünk¹²). Az η arányossági tényező beállításával meghatározhatjuk a tényleges eseményszámot.

A tesztjeinkben a $4\pi\eta a^2$ értéket felfelé kerekítettük és annak megfelelő számú téglalapot generáltunk (futásonként és adatsoronként).

A futtatásainkhoz a következő arányossági tényező értékeket (azaz lefedettségi arányokat) választottuk:

- $\eta = 1/100$, ehhez ≈ 10 téglalap,
- $\eta = 1/60$, ehhez ≈ 20 téglalap,

¹² Megjegyezzük, hogy viszonylag magas eseményszám esetén az átfedésektől nem tekinthetünk el; viszont valószínű, hogy a valós Omega kimenet (4.7. ábra) esetében eltekinthetünk az átfedések okozta extra területjárulék figyelembevételétől.

- $\eta = 1/20$, ehhez ≈ 60 téglalap,
- $\eta = 1/6$, ehhez ≈ 200 téglalap tartozik.

Megjegyezzük, hogy minden esetben törekedtünk arra, hogy a szimulált adatsoraink minél több jellemzőjét mutassák az Omega Pipeline feldolgozó eljárás kimeneti fájljainak. Az Omega infrastruktúra kimeneti fájljai – az utófeldolgozást követően – jó közelítéssel átlagosan 10-100 elemi eseményjelöltet (*tile*-t) tartalmaznak. Itt jegyezzük meg, hogy az adatsorok által lefedett időintervallum megállapításánál figyelembe kell venni a detektorok közötti távolságot¹³ (a véges, c sebességű GW-jelterjedés miatt).

Továbbá a valódi Omega Pipeline kimeneti adatoknál a téglalapok pozícióinak, illetve méreteinek (\sim szórásnak) idő és frekvencia szerinti pontos eloszlása egyelőre nem ismert, arra vonatkozóan csak becsléseink vannak. Elméleti modell híján tesztjeink során az egyenes eloszlást választottuk. Viszont ismert, hogy a valódi detektor zaj speciális eloszlást követ[14], az viszont kérdéses, mennyire torzítja a szimulációs eredményeinket a nem LIGO-specifikus zajeloszlás alkalmazása.

Itt jegyeznénk meg, hogy a jövőben, programunk hivatalos, többlépcsős ellenőriztetési folyamata során, illetve után lehetőségünk lesz hozzáférni nagy mennyiségű valódi Omega adatsorhoz, amelyek alapján a fenti statisztikák elkészíthetők, így algoritmusunk tesztelését még valóságosabb feltételek mellett folytatjuk.

A következőkben közöljük a szimulációkban kapott eredményeinket.

A fenti, kitüntetett η értékeken kívül elvégeztük az algoritmusunk tesztelését más értékekre is, kisebb pontossággal. Megvizsgáltuk a szimulációból származó teszteredményeink (elsősorban a normált CC értékek) eloszlását, hisztogramokat készítettünk bizonyos eseményszámok esetén¹⁴ (4.8. ábra). Az eredményeink azt mutatják, hogy a fentebb ismertetett tesztelési környezetben – 3-nál magasabb és 300-nál alacsonyabb számú esemény esetén a

¹³ A távolság alapján az egymással potenciálisan korreláló jelek közötti maximális detektálási-idő különbségre felső becslést adhatunk (a LIGO detektorok esetében ez az érték 10 ms). Nyilvánvaló, hogy a két adatsor keresztkorrelációs módszerekkel történő összehasonlítása ezen időkorlátot túl értelmét veszti, hiszen akkor korrelációt nem várhatunk. A GW-detektálási standardnak megfelelően minden detektor esetén az adatsorok tárolását definiált időintervallumokra feldarabolva végzik, ún. keretekben (*frame*-ekben). Ez technikailag azt jelenti, hogy a maximális korrelációs időértéket – bizonyos detektálási időintervallumban – összehasonlítható mintafájl darabszámmá konvertálhatjuk.

¹⁴ A kitöltési tényező, vagy lefedettség arány értéke nagyban függhet az események közötti, akár többszörös átfedésektől magasabb elemi eseményszám mellett; ezért a η faktor helyett a különböző esetekre a téglalapok számával utalunk.

normált hasonlóság mértéke, mint valószínűségi változó Poisson-eloszlást követő nagyszámú minta mellett. Azt találtuk, hogy a görbe maximuma a fentebb említett eseményszám intervallumban nagyrészt közel konstans CC értéknél található, félértékszélessége minden esetben ≈ 0.05 érték körüli. Viszont az eloszlás lecsengése függést mutat az események számától: a vizsgált magasabb eseményszámokra (≈ 100) meredekebb, mint ≈ 10 esemény esetén.

Megjegyezzük, hogy a 13, 21, 62 és 214 téglalapot generáló, 20 000 futtatást (összehasonlítást) magába foglaló tesztünkben kapott hisztogramok esetén az utolsó (legnagyobb) hasonlóság érték $CC = 0.48$.

Továbbá a 314, illetve annál több esemény összehasonlítását vizsgáló teszt sorozatunk eredményeit tanulmányozva azt találtuk, hogy – valószínűleg a sokszoros átfedések domináns jellege, vagy az előálló igen magas lefedettségi arány miatt – az illesztés nem konvergált. Ez azt jelenti, hogy vagy a felfelé ívelő pontsorozat esetében is érdemes megvizsgálunk, hogy az utolsó iterációban számított normált CC érték magasabb-e egy előre definiált p limitnél, és amennyiben nem, célszerű hasonlóságnak tekinteni az utolsó értéket; vagy további tesztek futtatásával érdemes megvizsgálni, hogy a $p = 0.5$ érték választás megfelelő-e (esetleg növelése indokolt lehet).

Viszont láttuk, hogy az Omega Pipeline kimeneteként várható eseményszámok mellett a programunk megbízhatóan működik.

A tesztprogramunk közvetlen kimenetét, és az azok alapján előállított hisztogramokat tanulmányozva arra a konklúzióra jutottunk, hogy az **Overlapper** algoritmusunk nagyságrendileg 10, illetve 100 eseményt tartalmazó idő-frekvencia diagramok esetén megbízhatóan alkalmazható, ugyanis:

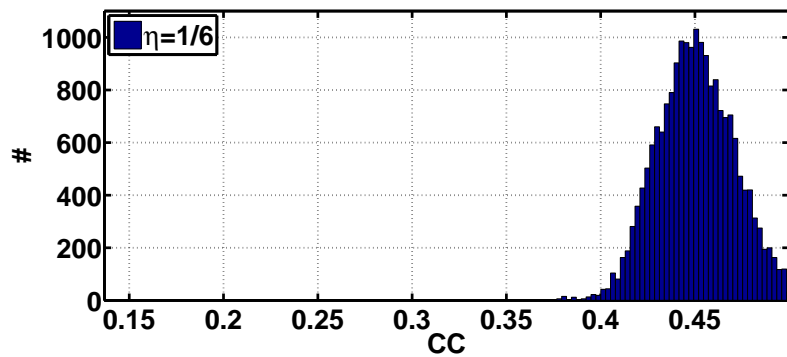
- megfelelően alacsony hasonlóság értéknél ($CC \approx 0.28$) jelentkezik tiszta zaj esetén az eloszlás maximuma;
- az eloszlás – folytonos határesetet tekintve – kellően gyorsan lecseng; minden esetben a fenti eseményszám tartományban értéke a $CC = 0.5$ kritikusnak tekintett normált hasonlóság értéknél a maximum közel ezred vagy annál kisebb részére csökken, illetve
- az eloszlás félértékszélessége megfelelően alacsony érték, továbbá igen jó közelítéssel állandó a fenti eseményszám intervallumban.

A tesztjeink alapján az egyes η értékek mellett a Λ **vágási határ**nak a következő értékeket választottuk:

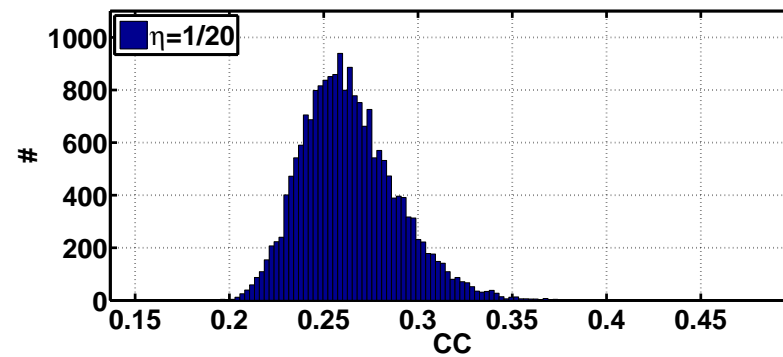
- $\eta = 1/100$ esetén $\Lambda = 0.30$
- $\eta = 1/60$ esetén $\Lambda = 0.28$
- $\eta = 1/20$ esetén $\Lambda = 0.31$
- $\eta = 1/6$ esetén $\Lambda = 0.48$.

Tehát egy adott lefedettségi arány esetén a fenti érték alatti hasonlósággal jellemezhető elemi eseményseregeket elvetjük, felette viszont riasztást adunk ki.

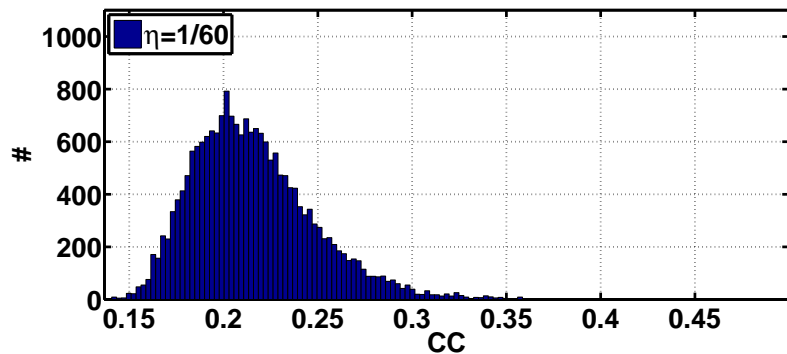
A vágási határnak azon normált hasonlóságot választottuk, amely értékig az egyes hisztogramokon az összehasonlítások 95%-a megtalálható (minden esetben összesen 20 000 minta esetén). Az így elért FAR érték 0.05.



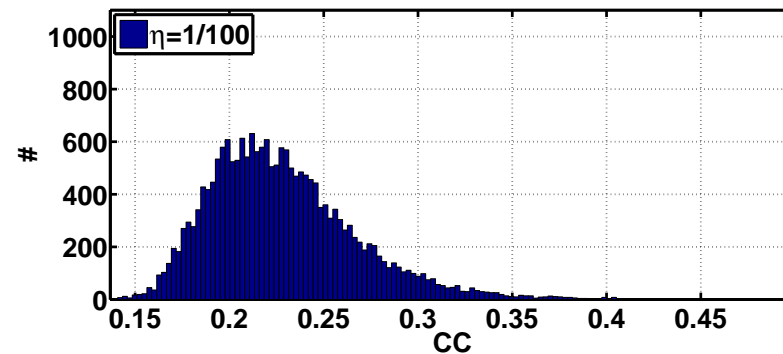
(a)



(b)



(c)



(d)

4.8. ábra. $\eta = 1/6, 1/20, 1/60, 1/100$ értékek mellett az események normált hasonlóságának eloszlása (minden esetben 20 000 minta esetén).

5. fejezet

Összefoglalás

A kitörés típusú gravitációshullám-jelek keresése ma a gravitációs hullámok kutatásának egyik fő irányvonalát képviseli. Az ilyen jeleket kibocsátó források becsült magas száma komoly lehetőséget és reményt ad az eddig csupán közvetett úton kimutatott gravitációs hullámok közvetlen detektálására.

A valós idejű kereséseknél széles körben alkalmazott, kitörések detektálására kifejlesztett Omega Pipeline inkoherens keresőeljárás utolsó lépésként összehasonlítja az egyes detektorok adatsoraiból előállított idő-frekvencia spektrogramokat. A téves riasztások magas száma mutatja, hogy az alkalmazott koincidencia analízis nem optimális.

Kifejlesztettünk egy algoritmust, ami egy vétó eljárás szerepét tölti be a keresés során. A program működése komplex alakzatok hasonlóságának gyors megállapításán alapszik. Az analízist egy optimalizált, keresztkorrelációs algoritmuson alapuló iterációs eljárás alkalmazásával végezzük el. Az eljárásunk azon a feltevésen alapszik, hogy asztrofizikai forrásból származó jel különböző detektorok adatának idő-frekvencia térbeli reprezentációjába hasonló alakzatokat létrehozva projektálódik.

Az algoritmusunkat teszteltük mind egyszerű alakzatok összehasonlításával, mind szimulált zajt tartalmazó adatsorok analízisével. A szimulációk alapján készítettünk teszt statisztikát, majd meghatároztuk az algoritmusunk téves riasztási arányát.

Az eredményeink azt mutatják, hogy az implementált algoritmusunk alkalmazásával számottevő mértékben csökkenthető a napjainkban gravitációs-hullám keresésre használt inkoherens eljárás téves riasztási valószínűsége. Továbbá a programunk kellően gyors a valós idejű analízishez. Ezek mellett az eljárásunk alkalmas alakzatfelismerésre is, amely révén lehetőségessé válhat részlegesen ismert jelalakok gyors keresése a zajos adatsorban.

A LSC-ben (*LIGO Scientific Collaboration*) való részvételünk a későbbiekben a LIGO adatsoraihoz való hozzáférést lehetővé teszi. Terveink között szerepel az algoritmusunk további fejlesztése immár valós adatokkal végzett tesztek segítségével, továbbá LIGO technikai, illetve tudományos cikk publikálása (a valós adat analízis során elért eredmények összefoglalásaként). Az algoritmusunkról LIGO konferencia poszter készült[24].

A jövőben integrálni tervezzük programunkat az Omega Pipeline infrastruktúrába, hogy azt valós gravitációs hullám keresések során is alkalmazhassuk. Ezt két szinten kívánjuk megvalósítani:

- vétó eljárásként a téves riasztások arányának csökkentésére;
- az Omega Scan-be: a legnagyobb energiájú trigger események gyors, automatikus összevetéséhez a PEM adataival.

Köszönetnyilvánítás

Köszönetet mondok témavezetőmnek, Raffai Péternek (ELTE), továbbá Frei Zsoltnak (ELTE), segítő és támogató munkájukért. A LIGO kollaboráció részéről köszönetet mondok Anthony Searle-nek és Jameson Rollinsnak, az Omega Pipeline fejlesztőinek, akik munkámat hasznos észrevételeikkel segítették. Végül köszönetet mondok az ELTE Bolyai Kollégium közösségének, Surján Péter igazgató úrnak az inspiráló szakmai környezet megteremtéséért.

Irodalomjegyzék

- [1] Hulse R. A., J. H. Taylor, *Astrophys. J.*, 195, L51-L53, 1975
- [2] Taylor J. H., *Rev. Mod. Phys.*, 66, 711, 1994
- [3] Weiss R., *Quarterly Progress Report of the Research Laboratory of Electronics*, 105, 54, 1972
- [4] <http://www.ligo.caltech.edu>
- [5] <http://www.minigrail.nl>
- [6] Thorne K. S., *300 Years of Gravitation*, ed. Hawking, S. W., Israel, W., Cambridge University Press, p. 330 – 458, 1987
- [7] <http://www.ligo.caltech.edu/advLIGO>
- [8] Shourov K. C., PhD tézis, MIT, 2005
- [9] Hughes S. A., Sz. Márka, P. L. Bender et al., *Proc. of the APS/DPF/DPB Summer Study on the Future of Particle Physics (Snowmass 2001)* ed. Graf, N. eConf, C010630, 402, astro-ph/0110349, 2001
- [10] Cutler C., K. S. Thorne, *Proceedings of 16th International Conference on General Relativity and Gravitation (Eds N.T. Bishop and S.D. Maharaj)*; gr-qc/0204090, 2002
- [11] Fryer C. L. et al., *Astrophys. J.*, 565, p. 430 – 446, 2002
- [12] Kotake K. et al., *Rept. Prog. Phys.*, 69, p. 971 – 1144, 2006
- [13] van Putten M. H. P. M., előadás, GRBs in the Afterglow Era: 3rd Workshop, Rome; Report No.: LIGO-P030003-00-R, 2002
- [14] Weiss R., *Quarterly Progress Report of RLE, MIT*, 105, p. 54, 1972

- [15] Sandberg V., S5 Report to the LSC, LIGO Scientific Collaboration Meeting, LIGO-060054-00-Z, 2006
- [16] http://www.ligo.caltech.edu/LIGO_web/other_gw/gw_projects.html
- [17] Sutton P. J. et al., X-Pipeline: An analysis package for autonomous gravitational-wave burst searches, New Journal of Physics, megjelenés alatt, arXiv:gr-qc/0908.3665v2, 2010
- [18] Klimenko S. et al., Coherent method for detection of gravitational wave bursts, Proceedings of Amaldi conference in Sydney, Australia, arXiv:gr-qc/0802.3232v2, 2008
- [19] <http://geco.phys.columbia.edu/omega>
- [20] Searle A. C., P. J. Sutton, M. Tinto, Bayesian detection of unmodeled bursts of gravitational waves, arXiv:gr-qc/0809.2809, 2009
- [21] Gabor D., Theory of Communication, Journal of Institution of Electrical Engineers, vol. 93, no. 26, p. 429 – 457, 1946
- [22] The LIGO Scientific Collaboration & The Virgo Collaboration, An upper limit on the stochastic gravitational-wave background of cosmological origin, Nature, Vol. 460, p. 990 – 994, 2009
- [23] Raffai P., MSc tézis, ELTE, 2006
- [24] Angler G., P. Raffai, A 2D Cross-correlational Veto Method for Incoherent Gravitational Wave Data Analysis Pipelines, LIGO Document G1000200, 2010
- [25] <https://ldas-jobs.ligo-la.caltech.edu/~detchar/S6/loudestOmega/calendar>